THE UNIVERSITY OF ALBERTA

MOMENT GENERATING FUNCTIONAL EQUATIONS OF CERTAIN STOCHASTIC LEARNING MODELS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

by

Yin Yew Hui

EDMONTON, ALBERTA

August, 1961

ABSTRACT

This thesis deals with the derivation of the moment generating functional equations of certain stochastic learning models and the solution of them in some special cases. The first chapter reviews the relevant materials in Bush and Mosteller's "Stochastic Models for Learning" [1]. The second chapter deals with the derivation of the asymptotic moment generating functional equations of the two experimenter-controlled events model and of the two subject-controlled events model. The third chapter deals with the solution of these equations in some special cases and the inversion of the solutions into distribution functions.

ACKNOWLEDGEMENT

I express here my thanks to Dr. John Ross McGregor for his suggestion of the problem and for his assistance in the preparation of this thesis. I also wish to express my appreciation to Dr. T. V. Narayana for his interest in the problem.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

SUMMARY OF THE RELEVANT LITERATURE ON STOCHASTIC LEARNING MODELS

In this introductory chapter, the work of Robert R. Bush and Frederick Mosteller (1) which is directly related to the present thesis is discussed. This chapter deals with the description and some of the mathematics of the Bush-Mosteller learning models of experimenter-controlled and subject-controlled events. It is divided into three sections. Section one gives the description and the mathematics common to both models; section two deals with experimenter-controlled events and section three deals with subject-controlled events.

Section One

The behavioural changes of learning are described here in terms of responses. In general, consider an experiment involving r mutually exclusive and exhaustive responses denoted by A_j where $j = 1, 2, \ldots r$. Each of these r responses has a certain probability of occurring and we represent each of these probabilities by p_j where $j = 1, 2, \ldots r$ respectively. Since the responses are exhaustive and mutually exclusive, the probability invariance rule holds, i.e.

$$(1.1.1) \qquad \sum_{j=1}^{r} p_j = 1 \qquad \text{and} \qquad 0 \leq p_j \leq 1 \ .$$

For two exhaustive and mutually exclusive responses,

$$(1.1.2) \qquad \sum_{j=1}^{2} p_j = p_1 + p_2 = 1 \quad .$$

The probabilities of the occurrences of the responses, however, do not necessarily remain constant; in the learning process, the occurrence of a certain event in the process will in general alter the probabilities of responses p_j in the next trial in a certain way. Let there be t exhaustive and mutually exclusive events in each trial of the process and let these events be denoted by E_i (where $i = 1, 2, \ldots t$) . In fact, we are assuming that every time a response is made, an outcome follows, and the outcome will alter the probabilities of the responses in a certain way.

To put the above in another way, a learning experiment consists of a sequence of trials, on each of which one and only one response occurs. Each response occurrence has an outcome which alters the set of probabilities of the responses on the next trial.

To describe the effects of the events on the set of probabilities of responses, we make use of the concept of an operator. Let the set of r probabilities be represented by a column vector $\underline{p} = (p_1\ p_2\ \ldots\ p_r)'$ with $\sum_{j=1}^{r} p_j = 1$, and define a matrix operator $\underline{T_i}$ for each of the t events where

$$(1.1.3) \quad \underline{T}_i = \begin{pmatrix} U_{11,i} & U_{12,i} & \cdots\cdots & U_{1r,i} \\ U_{21,i} & U_{22,i} & \cdots\cdots & U_{2r,i} \\ \vdots & \vdots & \cdots\cdots & \vdots \\ U_{r1,i} & U_{r2,i} & \cdots\cdots & U_{rr,i} \end{pmatrix} \quad \text{and } i = 1, 2, \ldots t .$$

When the matrix operator $\underline{T}_i$ is applied to the probability vector $\underline{p}$, a new probability vector

$$(1.1.4) \quad \underline{T}_i \, \underline{p} = \begin{pmatrix} U_{11,i} & U_{12,i} & \cdots\cdots & U_{1r,i} \\ U_{21,i} & U_{22,i} & \cdots\cdots & U_{2r,i} \\ \vdots & \vdots & \cdots\cdots & \vdots \\ U_{r1,i} & U_{r2,i} & \cdots\cdots & U_{rr,i} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_r \end{pmatrix}$$

$$= \begin{pmatrix} \sum_j U_{1j,i} \, p_j \\ \sum_j U_{2j,i} \, p_j \\ \vdots \\ \sum_j U_{rj,} \, p_j \end{pmatrix} \qquad \text{where } \begin{array}{l} i = 1, 2, \ldots t \\ j = 1, 2, \ldots r \end{array}$$

is obtained.

Since $\underline{T}_i \, \underline{p}$ is a new probability vector,

$$\sum_{k=1}^{r} \sum_{j=1}^{r} U_{kj,i} \, p_j = 1 \qquad \text{where } (i = 1, 2, \ldots t)$$

for all values of p_j .

Consider the case where $p_1 = 1$, then $p_j = 0$ for $j = 2, 3, \ldots r$, and so

$$\sum_{k=1}^{r} U_{k1} = 1 , \quad \text{i.e.} \quad U_{11} + U_{21} + \ldots + U_{r1} = 1$$

where the subscript i has been dropped for simplicity. Thus by letting $p_j = 1$ for $j = 1, 2, \ldots r$ respectively, it can be established that each column of the matrix $\underline{T}_i$ must sum to unity,

$$\text{(1.1.5)} \quad \text{i.e.} \quad \sum_{k=1}^{r} U_{kj} = 1 \quad \text{for} \quad j = 1, 2, \ldots r .$$

Since the r classes of responses are arbitrarily defined, it should be mathematically possible to combine two classes of responses into one and to arrive at the same result as if we started off with $r - 1$ classes. Consider the case where we want to combine classes one and two into a new class. This can be done by multiplying the vector $\underline{p}$ by the $r \times r$ matrix $\underline{c}$ where

$$\text{(1.1.6)} \quad \underline{c} = \begin{pmatrix} 1 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix} .$$

Since

$$\underline{T}_i\, \underline{p} = \begin{pmatrix} \sum_j U_{1j}\, p_j \\ \sum_j U_{2j}\, p_j \\ \sum_j U_{3j}\, p_j \\ \vdots \\ \sum_j U_{rj}\, p_j \end{pmatrix} \qquad \text{where } \; i = 1, 2, \dots t \quad j = 1, 2, \dots r ,$$

it follows that

$$(1.1.7) \quad \underline{c}\, \underline{T}_i\, \underline{p} = \begin{pmatrix} \sum_j (U_{1j} + U_{2j})\, p_j \\ 0 \\ \sum_j U_{3j}\, p_j \\ \vdots \\ \sum_j U_{rj}\, p_j \end{pmatrix} .$$

Since we require that classes one and two combine to form a new class with probability $p_c = p_1 + p_2$, the components of the vector $\underline{c}\, (\underline{T}_i\, \underline{p})$ must not depend on p_1 and p_2 individually but on $p_c = p_1 + p_2$, and so

$$\begin{aligned} U_{31} &= U_{32} = U_3 \\ U_{41} &= U_{42} = U_4 \\ &\vdots \\ U_{r1} &= U_{r2} = U_r , \qquad \text{say} . \end{aligned}$$

Since every column vector of $\underline{T}_i$ must sum to 1 ,

$$U_{11} = 1 - U_{21} - U_3 - U_4 \cdots - U_r \quad .$$

In a similar way, by combining classes one and three of the responses into one, it can be shown that

$$U_{21} = U_{23} = U_2 \text{, say.}$$

Thus

$$U_{11} = 1 - U_2 - U_3 \cdots - U_r \ ,$$

and the first column of $\underline{T}_i$ is

$$\begin{matrix} 1 - U_2 - U_3 \cdots - U_r \\ U_2 \\ U_3 \\ \cdot \\ \cdot \\ \cdot \\ U_r \end{matrix} \quad .$$

Similarly by combining classes two and three we obtain

$$U_{12} = U_{13} = U_1 \text{ , say,}$$

and thus the second column of $\underline{T}_i$ is

$$\begin{matrix} U_1 \\ 1 - U_1 - U_3 - U_4 \cdots - U_r \\ U_3 \\ \cdot \\ \cdot \\ \cdot \\ U_r \end{matrix} \qquad .$$

By using this method it can be shown that

$$(1.1.8) \quad \underline{T}_i = \begin{pmatrix} 1-U_2-U_3\cdots-U_r & U_1 & \cdots & U_1 \\ U_2 & 1-U_1-U_3-U_4\cdots-U_r & \cdots & U_2 \\ U_3 & U_3 & \cdots & U_3 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ U_r & U_r & \cdots & 1-U_1-U_2\cdots-U_{r-1} \end{pmatrix}$$

$$= (1 - U_1 - U_2 - \cdots - U_r) \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} + \begin{pmatrix} U_1 & U_1 & \cdots & U_1 \\ U_2 & U_2 & \cdots & U_2 \\ U_3 & U_3 & \cdots & U_3 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ U_r & U_r & \cdots & U_r \end{pmatrix}$$

Let $\alpha_i = 1 - U_1 - U_2 - U_3 - \dots - U_r = 1 - \sum_{j=1}^{r} U_j$,

$U_j = (1 - \alpha_i)\lambda_j$, where $j = 1, 2, \dots r$.

Then

$$\underline{T}_i = \alpha_i \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 \end{pmatrix} + \begin{pmatrix} (1-\alpha_i)\lambda_1 & (1-\alpha_i)\lambda_1 & \dots & (1-\alpha_i)\lambda_1 \\ (1-\alpha_i)\lambda_2 & (1-\alpha_i)\lambda_2 & \dots & (1-\alpha_i)\lambda_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ (1-\alpha_i)\lambda_r & (1-\alpha_i)\lambda_r & \dots & (1-\alpha_i)\lambda_r \end{pmatrix} \tag{1.1.9}$$

and if we put

$$\underline{\Lambda}_i = \begin{pmatrix} \lambda_1 & \lambda_1 & \dots & \lambda_1 \\ \lambda_2 & \lambda_2 & \dots & \lambda_2 \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \lambda_r & \lambda_r & \dots & \lambda_r \end{pmatrix} , \tag{1.1.10}$$

then

$$\underline{T}_i = \alpha_i \underline{I} + (1 - \alpha_i)\underline{\Lambda}_i \tag{1.1.11}$$

and

$$\underline{\Lambda}_i \underline{p} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \cdot \\ \cdot \\ \lambda_r \end{pmatrix} = \underline{\lambda}_i \quad . \tag{1.1.12}$$

Thus

(1.1.13) $$\underline{T}_i\ \underline{p} = \alpha_i\ \underline{I}\ \underline{p} + (1 - \alpha_i)\underline{\Lambda}_i\ \underline{p}$$

$$= \alpha_i\ \underline{p} + (1 - \alpha_i)\underline{\lambda}_i$$

Therefore, for the case $r = 2$, $i = 1, 2 \dots t$,

$$\underline{T}_i = \begin{pmatrix} 1 - U_{2i} & U_{1i} \\ U_{2i} & 1 - U_{1i} \end{pmatrix} .$$

Let $U_{2i} = b_i$, $U_{1i} = a_i$, $\underline{p} = \begin{pmatrix} p \\ q \end{pmatrix}$,

where $p + q = 1$, and so $\alpha_i = 1 - a_i - b_i$.

Then

(1.1.14) $$\underline{T}_i = \begin{pmatrix} 1 - b_i & a_i \\ b_i & 1 - a_i \end{pmatrix} , \qquad \text{and}$$

(1.1.15) $$\underline{T}_i\ \underline{p} = \begin{pmatrix} (1 - b_i)p + a_i q \\ b_i p + (1 - a_i)q \end{pmatrix} = \begin{pmatrix} Q_i\ p \\ \tilde{Q}_i\ q \end{pmatrix}$$

where $Q_i\ p = (1 - b_i)p + a_i\ q$

$\tilde{Q}_i\ q = b_i\ p + (1 - a_i)q \qquad i = 1, 2 \dots t$.

By rearranging the terms on the right hand side, the following forms for $Q_i\ p$ are obtained

(a) SLOPE-INTERCEPT FORM:

(1.1.16) $Q_i\ p = a_i + \alpha_i\ p \qquad$ where $\qquad i = 1 , 2 \dots t,$

(b) GAIN-LOSS FORM:

(1.1.17) $Q_i\ p = p + a_i(1 - p) - b_i\ p$ where $i = 1, 2, \ldots t$

(c) FIXED-POINT FORM:

(1.1.18) $Q_i\ p = \alpha_i\ p + (1 - \alpha_i)\lambda_{1i}$ where $i = 1, 2 \ldots t$

Since $\tilde{Q}_i\ q = 1 - Q_i\ p$, by the fixed-point form,

$$\tilde{Q}_i\ q = \alpha_i\ q + (1 - \alpha_i)(1 - \lambda_{1i}) \quad .$$

Dropping the subscript 1 in λ_{1i}, we have

(1.1.19)
$$\underline{T_i}\ \underline{p} = \begin{pmatrix} \alpha_i\ p \\ \alpha_i\ q \end{pmatrix} \begin{matrix} + \\ + \end{matrix} \begin{pmatrix} (1 - \alpha_i)\lambda_i \\ (1 - \alpha_i)(1 - \lambda_i) \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_i\ p \\ \alpha_i\ q \end{pmatrix} + \begin{pmatrix} (1 - \alpha_i)\lambda_i \\ (1 - \alpha_i)(1 - \lambda_i) \end{pmatrix}$$

$$= \alpha_i \begin{pmatrix} p \\ q \end{pmatrix} + (1 - \alpha_i) \begin{pmatrix} \lambda_i \\ 1 - \lambda_i \end{pmatrix}$$

$$= \alpha_i\ \underline{p} + (1 - \alpha_i)\underline{\lambda_i} \quad .$$

As $Q_i\ p$ and p both have to satisfy the relations $0 \leq Q_i\ p \leq 1$ $0 \leq p \leq 1$, the parameters a_i, b_i and α_i have to obey certain restrictions.

Consider the gain-loss form,

$$Q_i\, p = p + a_i(1 - p) - b_i\, p \ .$$

Since $0 \leq Q_i\, p \leq 1$ it follows that

(1.1.20) $0 \leq p + a_i(1 - p) - b_i\, p \leq 1$.

Let $p = 0$, then the relation

(1.1.21) $0 \leq Q_i\, p \leq 1$ becomes $0 \leq a_i \leq 1$.

For $p = 1$, the relation $0 \leq Q_i\, p \leq 1$ becomes

(1.1.22) $0 \leq 1 - b_i \leq 1$ or $1 \geq b_i \geq 0$.

Hence the admissible region for a_i and b_i is shown by the shaded part of figure 1.1 .

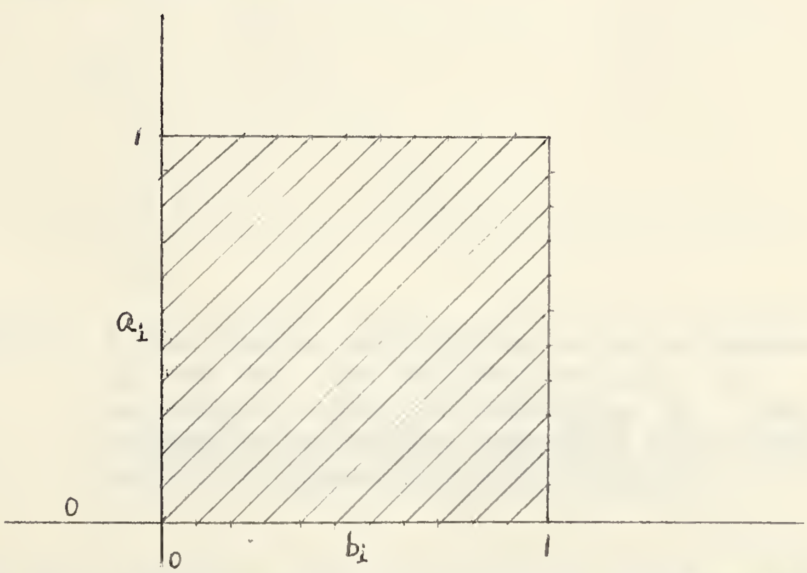

Fig. 1.1 A plot showing restrictions on the parameters a_i and b_i . The square shaded area represents the admissible region. (Reproduced from [1] page 34)

Since $0 \leq a_i \leq 1$ and $0 \leq p_i \leq 1$, the parameter α_i, which is equal to $1 - a_i - b_i$, must then obey the relation $-1 \leq \alpha_i \leq 1$. However, this restriction is not strong enough to keep $Q_i\ p = a_i + \alpha_i\ p$ in the closed interval $[0,1]$. For instance, if $a_i = 0$, and $\alpha_i = -1$, then $Q_i\ p = -p$ which is impermissible. This shows that a_i and α_i are not independent. For a positive α_i, $Q_i\ p = a_i + \alpha_i\ p$ is a maximum when $p = 1$, i.e. $a_i + \alpha_i \leq 1$ or $\alpha_i \leq 1 - a_i$. Therefore α_i must lie in the shaded area of the following figure.

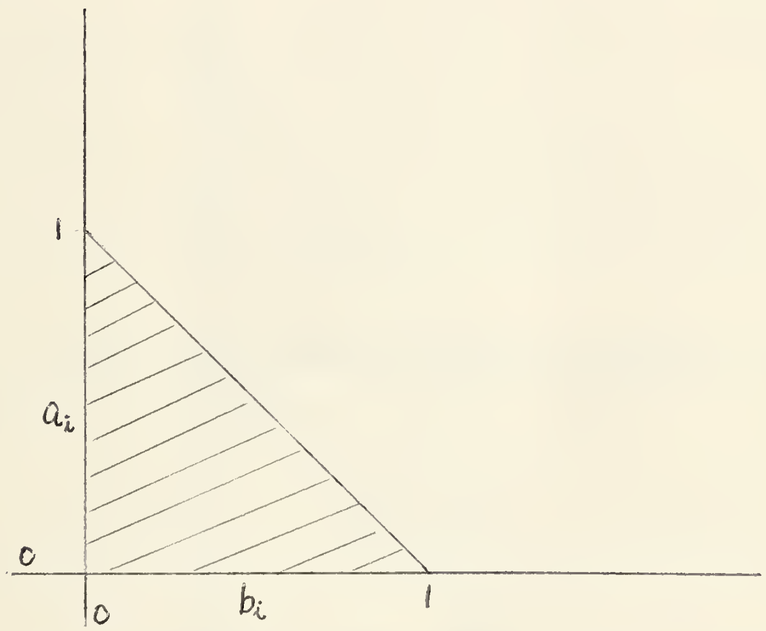

Fig. 1.2 A plot showing restrictions on the parameters a_i and b_i when the function $Q_i\ p$ is restricted to have positive slope α_i. The triangular shaded area represents the admissible region. (Reproduced from [1] page 34)

When α_i is negative, $Q_i\ p = a_i + \alpha_i\ p$ becomes a maximum when $p = 0$ and so $a_i \leq 1$; $Q_i\ p = a_i + \alpha_i\ p$ becomes minimum if $p = 1$.

i.e. $\quad 0 \leq a_i + \alpha_i$

(1.1.23) $\quad -a_i \leq \alpha_i$.

Thus by combining the above inequalities for positive and negative α's , we have

(1.1.24) $\quad -a_i \leq \alpha_i \leq 1 - a_i$.

This range of possible values of a_i and α_i is shown in the following figure.

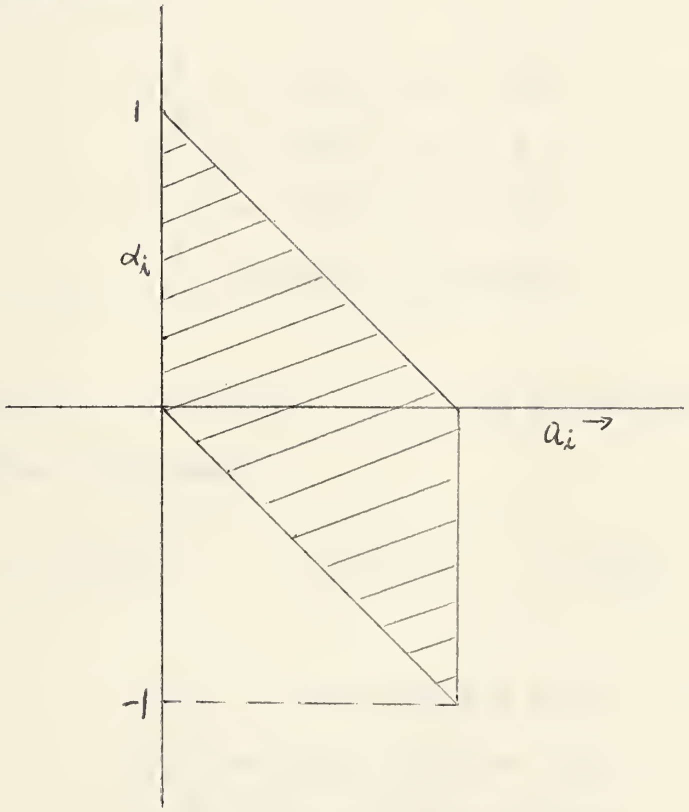

Fig. 1.3 Plot showing the possible values of α_i and a_i . The shaded area indicates the values consistent with the restrictions. $0 \leq a_i \leq 1$, $-a_i \leq \alpha_i \leq 1 - a_i$. (Reproduced from [1] page 35)

For the fixed-point form

$$Q_i\, p = \alpha_i\, p + (1 - \alpha_i)\lambda_i \quad ,$$

it follows from the above discussion that the restriction is

(1.1.25) $0 \leq \lambda_i \leq 1$.

When α_i is positive, $Q_i\, p$ has an increasing type of character. For example,

let $\alpha_i = 0.5 \qquad \lambda_i = 1 \qquad p = 0.2$.

Then
$$\begin{aligned} Q_i^{\,1}\, p &= 0.10 + 0.5 = 0.60 \\ Q_i^{\,2}\, p &= 0.30 + 0.5 = 0.8 \\ Q_i^{\,3}\, p &= 0.40 + 0.5 = 0.90 \\ Q_i^{\,4}\, p &= 0.45 + 0.5 = 0.95 \quad . \end{aligned}$$

However, when α is negative, $Q_i\, p$ has an oscillating type of character. For example

let $\alpha_i = -0.5 \qquad \lambda_i = 0.5 \qquad p = 0.2$.

Then
$$\begin{aligned} Q_i^{\,1}\, p &= -0.10 \;\; + 0.75 = 0.65 \\ Q_i^{\,2}\, p &= -0.325 + 0.75 = 0.425 \\ Q_i^{\,3}\, p &= -0.213 + 0.75 = 0.538 \\ Q_i^{\,4}\, p &= -0.269 + 0.75 = 0.481 \quad . \end{aligned}$$

Throughout the rest of this thesis, we shall restrict ourselves to $0 \leq \alpha_i$. Thus the restrictions on the parameters are as follows:

(a) SLOPE-INTERCEPT FORM:

(1.1.26) $Q_i\, p = a_i + \alpha_i\, p$

$$0 \le a_i \le 1 \qquad\qquad 0 \le \alpha_i \le 1 - a_i$$

(b) GAIN-LOSS FORM:

(1.1.27) $Q_i\, p = p + a_i(1 - p) - b_i\, p$

$$0 \le a_i \le 1 \qquad\qquad 0 \le b_i \le 1 \qquad\qquad 0 \le a_i + b_i \le 1$$

(c) FIXED-POINT FORM:

(1.1.28) $Q_i\, p = \alpha_i\, p + (1 - \alpha_i)\lambda_i$

$$0 \le \lambda_i \le 1 \qquad\qquad 0 \le \alpha_i \le 1 \quad .$$

Repetitive Application of the Operator $\underline{T}_i$

As shown above,

$$\underline{T}_i = \alpha_i\, \underline{I} + (1 - \alpha_i)\underline{\Lambda}_i \ , \quad \text{and so}$$

$$\underline{T}_i^2 = \underline{T}_i \cdot \underline{T}_i = \left[\alpha_i\, \underline{I} + (1 - \alpha_i)\underline{\Lambda}_i\right] \cdot \left[\alpha_i\, \underline{I} + (1 - \alpha_i)\underline{\Lambda}_i\right]$$

$$= \alpha_i^2\, \underline{I}^2 + \alpha_i(1 - \alpha_i)\underline{I}\,\underline{\Lambda}_i + \alpha_i(1 - \alpha_i)\underline{\Lambda}_i\,\underline{I}$$

$$+ (1 - \alpha_i)^2\underline{\Lambda}_i^2$$

However,

$$\underline{\Lambda}_i^2 = \begin{pmatrix} \lambda_{i1} \sum_j \lambda_{ij} & \lambda_{i1} \sum_j \lambda_{ij} & \cdots & \lambda_{i1} \sum_j \lambda_{ij} \\ \lambda_{i2} \sum_j \lambda_{ij} & \lambda_{i2} \sum_j \lambda_{ij} & \cdots & \lambda_{i2} \sum_j \lambda_{ij} \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_{ir} \sum_j \lambda_{ij} & \lambda_{ir} \sum_j \lambda_{ij} & \cdots & \lambda_{ir} \sum_j \lambda_{ij} \end{pmatrix} .$$

Since $\sum_{j=1}^{r} \lambda_{ij} = 1$,

$$\underline{\Lambda}_i^2 = \underline{\Lambda}_i \quad ,$$

and so

$$\underline{T}_i^2 = \alpha_i^2 \, \underline{I} + 2\alpha_i(1 - \alpha_i)\underline{\Lambda}_i + (1 - \alpha_i)^2 \underline{\Lambda}_i$$

$$= \alpha_i^2 \, \underline{I} + (1 - \alpha_i^2)\underline{\Lambda}_i \quad .$$

To prove that in general

(1.1.29) $\underline{T}_i^n = \alpha_i^n \, \underline{I} + (1 - \alpha_i^n)\underline{\Lambda}_i$,

assume that this holds for the nth case

i.e. $\underline{T}_i^n = \alpha_i^n \, \underline{I} + (1 - \alpha_i^n)\underline{\Lambda}_i$.

Then for the $(n + 1)$th case

(1.1.30) $\underline{T}_i^{n+1} = \underline{T}_i(\underline{T}_i^n) = \left[\alpha_i \, \underline{I} + (1 - \alpha_i)\underline{\Lambda}_i\right] \left[\alpha_i^n \, \underline{I} + (1 - \alpha_i^n)\underline{\Lambda}_i\right]$

$$= \alpha_i^{n+1} \, \underline{I} + (1 - \alpha_i)\alpha_i^n \underline{\Lambda}_i + \alpha_i(1 - \alpha_i^n)\underline{\Lambda}_i$$

$$+ (1 - \alpha_i)(1 - \alpha_i^n)\underline{\Lambda}_i^2$$

$$= \alpha_i^{n+1}\,\underline{I} + [(1-\alpha_i)\alpha_i^{n} + (1-\alpha_i^{n})\alpha_i + (1-\alpha_i)(1-\alpha_i^{n})]\underline{\Lambda}_i$$

$$= \alpha_i^{n+1}\,\underline{I} + (1-\alpha_i^{n+1})\,\underline{\Lambda}_i \quad ,$$

which is of the same form as in the nth case. Hence in general, (1.1.29) is true

i.e. $\quad \underline{T}_i^{\,n} = \alpha_i^{n}\,\underline{I} + (1-\alpha_i^{n})\underline{\Lambda}_i \quad .$

Thus by (1.1.12), applying $\underline{T}_i^{\,n}$ on $\underline{p}$, we have

(1.1.31) $\underline{T}_i^{\,n}\,\underline{p} = \alpha_i^{n}\,\underline{p} + (1-\alpha_i^{n})\underline{\lambda}_i \quad .$

Since $|\alpha_i| \leq 1$, for $|\alpha_i| < 1$,

$$\lim_{n\to\infty} \underline{T}_i^{\,n}\,\underline{p} = \lim_{n\to\infty} \alpha_i^{n}\,\underline{p} + (1-\alpha_i^{n})\underline{\lambda}_i = \underline{\lambda}_i \quad .$$

Hence for $|\alpha_i| < 1$, $\underline{\lambda}_i$ is the limiting vector of the application of the operator $\underline{T}_i$ on the vector $\underline{p}$ an infinite number of times. The following figure shows the effect of successive application of the operator Q_i on p when $p = 0.1$, $\lambda_i = 0.8$ and $\alpha_i = 0.9$.

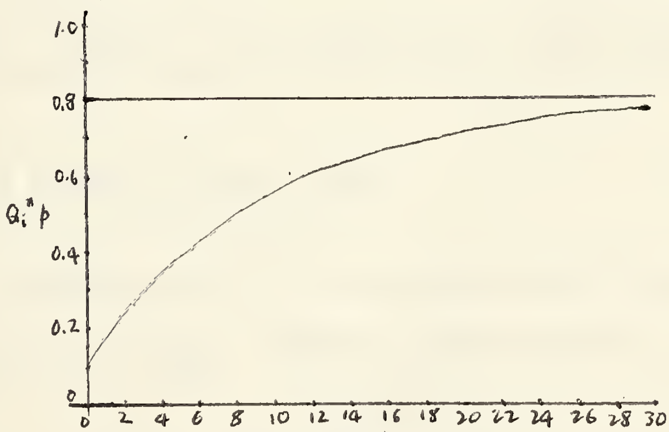

Fig. 1.4 Values of $Q_i^{\,n}\,p$ plotted against trial number n . $Q_i^{\,n}p = \alpha_i^{\,n}p + (1-\alpha_i^{\,n})\lambda_i$ with $p = 0.1$, $\lambda_i = 0.8$, and $\alpha_i = 0.9$ were used in plotting the curve. (Reproduced from [1] page 60)

In the case $r = 2$, $\underline{p} = \begin{pmatrix} p \\ 1 - p \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$,

$$(1.1.32) \quad \underline{T}_i^{\,n}\, \underline{p} = \underline{T}_i^{\,n} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} Q_i^{\,n}\, p \\ \tilde{Q}_i^{\,n}\, q \end{pmatrix} = \alpha_i^{\,n}\, \underline{p} + (1 - \alpha_i^{\,n})\, \underline{\lambda}_i$$

$$= \alpha_i^{\,n} \begin{pmatrix} p \\ q \end{pmatrix} + (1 - \alpha_i^{\,n}) \begin{pmatrix} \lambda_{1i} \\ \lambda_{2i} \end{pmatrix},$$

i.e. $\quad Q_i^{\,n}\, p = \alpha_i^{\,n}\, p + (1 - \alpha_i^{\,n})\lambda_{1i}$

$\quad \tilde{Q}_i^{\,n}\, q = \alpha_i^{\,n}\, q + (1 - \alpha_i^{\,n})\lambda_{2i}$.

Thus

$$(1.1.33) \quad \lim_{n \to \infty} Q_i^{\,n}\, p = \lim_{n \to \infty} \alpha_i^{\,n}\, p + (1 - \alpha_i^{\,n})\lambda_{1i}$$

$$= \lambda_{1i} \qquad \text{for} \qquad |\alpha_i| < 1 ,$$

$$(1.1.34) \quad \lim_{n \to \infty} \tilde{Q}_i^{\,n}\, q = \lim_{n \to \infty} \alpha_i^{\,n}\, q + (1 - \alpha_i^{\,n})\lambda_{2i}$$

$$= \lambda_{2i} .$$

We call λ_{1i} and λ_{2i} the limit points of the operator $\underline{T}_i$.

Concept of Symmetry for Two Events

Consider a case of two responses and two events with operators Q_1 and Q_2 . With the same notation as before,

$$Q_1\, p = \alpha_1\, p + (1 - \alpha_1)\lambda_1$$

$$Q_2\,p = \alpha_2\,p + (1 - \alpha_2)\lambda_2 \quad .$$

Let (1.1.35) $\widetilde{Q}_2\,q = 1 - Q_2\,p$

$$= 1 - \alpha_2\,p - (1 - \alpha_2)\lambda_2$$

$$= \alpha_2\,q + (1 - \alpha_2)(1 - \lambda_2) \quad ,$$

and define the symmetry requirement as follows:

$$Q_1\,p = \widetilde{Q}_2\,p \qquad \text{and} \qquad Q_1\,q = \widetilde{Q}_2\,q \quad .$$

This is true if and only if

(1.1.36) $\qquad \alpha_2 = \alpha_1$

(1.1.37) $\quad 1 - \lambda_2 = \lambda_1 \quad .$

The first condition states that the slope parameters are equal. This is called the equal α condition. The second condition requires that Q_1 and Q_2 have complementary limit points. The equal α-condition implies that event E_1 and event E_2 have "equal but opposite" effects on behaviour. E_1 has the same effect on response A_1 as E_2 has on response A_2. For experimenter-controlled events, the equal α condition gives minor simplification.

Section Two

Experimenter-Controlled Events

Events are defined to be experimenter-controlled if the events occur with probabilities fixed by the experimenter.

The case of two experimenter-controlled events involves two events on each trial of the experiment. For example, if in an experiment, there is only one response and if the reward of that response is taken to be an event E_1 and the non-reward of the response is taken to be event E_2 and if the proportions of reward and non-reward are fixed by the experimenter, then the experiment is two experimenter-controlled with one response.

Brunswik T-Maze Experiment

An example is the Brunswik T-Maze experiment which deals with the reward training of rats. Suppose there is an elevated T-Maze with boxes at the two ends of the T . A hungry rat is placed at the base of the T and is permitted to go along the Maze to the turning point where it can go either to the left or to the right. On each trial, some food is either placed in the box at the left or in the box at the right. The following assumptions are made.

(1) A right turn with a reward and a left turn without a reward produce the same effect on the probability p of turning right on the next trial and they constitute a single event E_1 with operator Q_1 .

(2) A right turn without reward and a left turn with reward have the same effect on p , the probability of turning right on the next trial, and they constitute another single event E_2 with operator Q_2 .

(3) During the trials, a proportion π_1 of the rewards are placed on the right side and a proportion $\pi_2 = 1 - \pi_1$ of the rewards are placed on the left.

The following are the four possible response-outcome pairs: --

Response	Outcome	Operator	Prob. of Occurrence
right turn	reward	Q_1	$p\,\pi_1$
left turn	non-reward	Q_1	$q\,\pi_1 = (1 - p)\pi_1$
right turn	non-reward	Q_2	$p(1 - \pi_1) = p\,\pi_2$
left turn	reward	Q_2	$q\,\pi_2 = (1 - p)\pi_2$.

From the above, the probability of the event E_1 is $p\,\pi_1 + (1 - p)\pi_1 = \pi_1$ and the probability of the event E_2 is $p\,\pi_2 + (1 - p)\pi_2 = \pi_2$, with $\pi_1 + \pi_2 = 1$. Since π_1 and π_2 are fixed by the experimenter and there are only two possible events on each trial, this is called the case of two experimenter-controlled events.

Consider the effects of applying operator Q_1 with fixed probability π_1 and operator Q_2 with fixed probability $\pi_2 = 1 - \pi_1$ in the case of two experimenter-controlled events with responses A_1 and A_2 . Let the experiment be run through by a large number of initially identical animals or organisms for n trials. Since there are two possible events for each trial, and there are n trials, the total number of possible sequences or groups is 2^n .

Let $p_{\nu n}$ be the probability of response A_1 of the νth group on the nth trial and let $P_{\nu n}$ be the proportion of the population in the νth group. Then $\bar{p}_n$, the average probability of the response A_1 for the whole population on the nth trial is

$$(1.2.1) \quad \bar{p}_n = \sum_{\nu=1}^{2^n} p_{\nu n} P_{\nu n} .$$

On the $(n+1)$th trial, the new values of p would be as follows:

New Values of p	New Proportion
$Q_1 p_{\nu n} = a_1 + \alpha_1 p_{\nu n}$	$\pi_1 P_{\nu n}$
$Q_2 p_{\nu n} = a_2 + \alpha_2 p_{\nu n}$	$\pi_2 P_{\nu n}$

and so,

$$(1.2.2) \quad \bar{p}_{n+1} = \sum_{\nu=1}^{2^n} \left[\pi_1 P_{\nu n}(a_1 + \alpha_1 p_{\nu n}) + \pi_2 P_{\nu n}(a_2 + \alpha_2 p_{\nu n}) \right]$$

$$= \sum_{i=1}^{2} \pi_i a_i + \sum_{i=1}^{2} (\pi_i \alpha_i)\bar{p}_n$$

$$= \bar{a} + \bar{\alpha}\,\bar{p}_n$$

where $\bar{a} = \sum_{i=1}^{2} \pi_i a_i = \pi_1 a_1 + \pi_2 a_2$; $\bar{\alpha} = \sum_{i=1}^{2} (\pi_i \alpha_i) = \pi_1 \alpha_1 + \pi_2 \alpha_2$.

Hence from the above a new operator $\bar{Q}$ is introduced such that

$$(1.2.3) \quad \bar{p}_{n+1} = \bar{Q}\,\bar{p}_n = \bar{a} + \bar{\alpha}\,\bar{p}_n .$$

Then by (1.1.31) the solution of the difference equation is

$$(1.2.4)\qquad \bar{p}_n = \bar{\alpha}^n p + (1 - \bar{\alpha}^n)\,\bar{\lambda} \qquad \text{where} \quad \bar{\lambda} = \frac{\bar{a}}{1 - \bar{\alpha}} \quad .$$

The following tree diagram shows the sequence of events together with their probabilities for the first three applications of the operators.

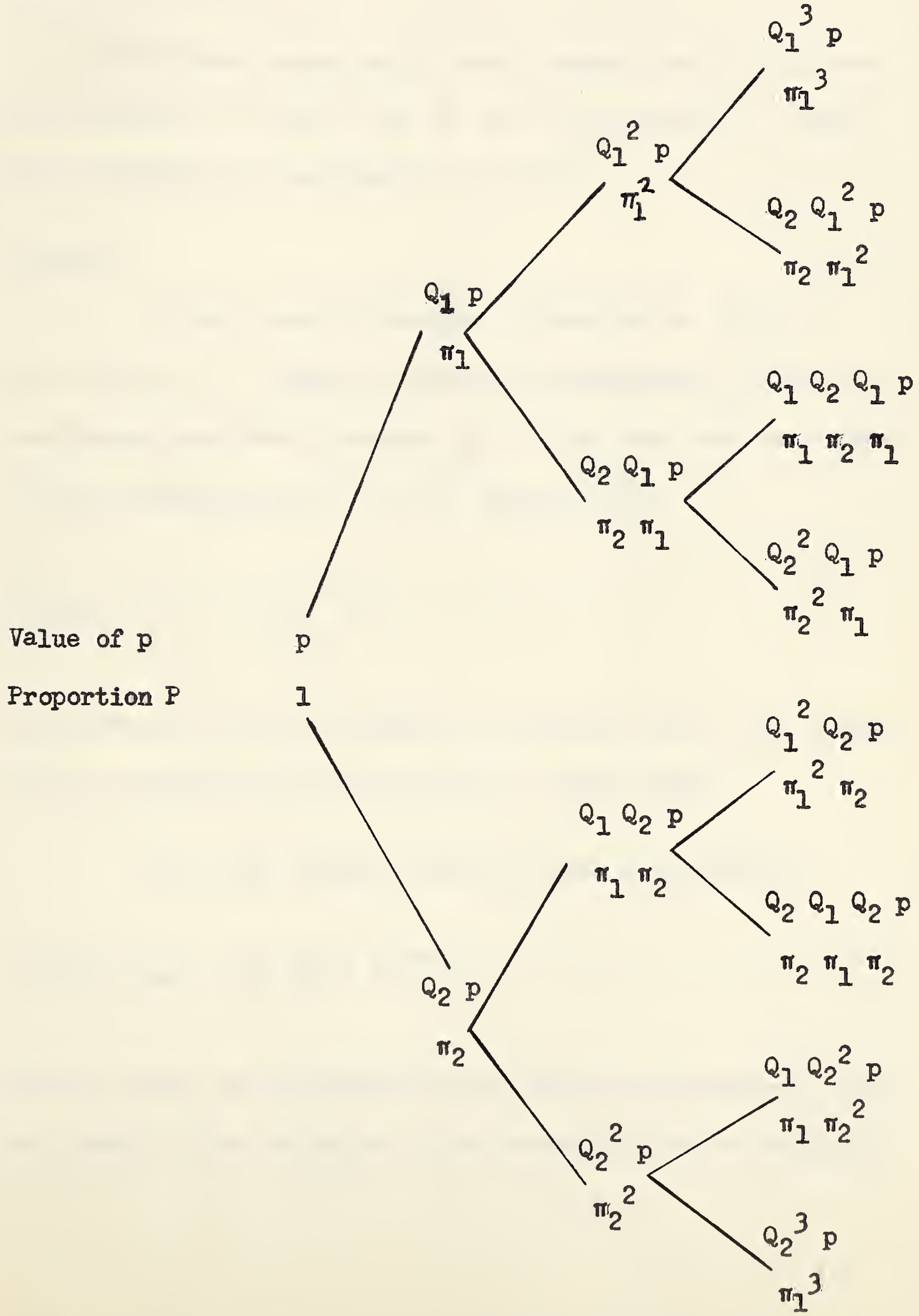

Fig. 1.5 The successive splits that a large group of animals

goes through after three events when the probability of applying operator Q_1 is π_1, given an operator is to be applied (experimenter-controlled events). The proportions in the groups written beneath the p value of the group have been written to parallel the Q's rather than in the simplest form -- thus under $Q_2 Q_1 Q_2 p$ is written $\pi_2 \pi_1 \pi_2$ rather than $\pi_1 \pi_2^2$. (Reproduced from (1) page 70)

It is clear upon examining the tree diagram that if Q_1 has been applied u times and Q_2 has been applied v times, the proportion in the group is $\pi_1^u \pi_2^v$.

Moments

Given a set of discrete probabilities $P_\nu \geq 0$, $\nu = 1, 2, \ldots$, finite or infinite in number and $\sum_\nu P_\nu = 1$, each associated with a number p_ν, then the mth raw moment of the distribution of p's is defined to be

$$(1.2.5) \qquad V_m = \sum_\nu p_\nu^m P_\nu .$$

In words, the mth raw moment is the mean of the mth power of the variable whose distribution is under study.

The mth moment about the mean is defined as

$$(1.2.6) \qquad \mu_m = \sum_\nu (p_\nu - V_1)^m P_\nu .$$

That is, the mth moment about the mean is the average of the mth power of the deviation of the variable from its own mean.

Moments for Two Experimenter-Controlled Events

On the $(n+1)$th trial, the mth raw moment of the two experimenter-controlled events distribution is given by the definition of moments to be

$$(1.2.7)\quad V_{m,n+1} = \sum_{i=1}^{2} \pi_i \left(\sum_{\nu=1}^{2^n} (Q_i\, p_{\nu n})^m P_{\nu n} \right)$$

$$= \pi_1 \sum_{\nu=1}^{2^n} (a_1 + \alpha_1\, p_{\nu n})^m P_{\nu n} + \pi_2 \sum_{\nu=1}^{2^n} (a_2 + \alpha_2\, p_{\nu n})^m P_{\nu n}$$

$$= \pi_1 \sum_{\nu} \sum_{u} \binom{m}{u} a_1^{m-u}\, \alpha_1^{u}\, p_{\nu n}^{u}\, P_{\nu n} + \pi_2 \sum_{\nu} \sum_{u} \binom{m}{u} a_2^{m-u}\, \alpha_2^{u}\, p_{\nu n}^{u}\, P_{\nu n}$$

$$= \pi_1 \sum_{u} \binom{m}{u} a_1^{m-u}\, \alpha_1^{u}\, V_{u,n} + \pi_2 \sum_{u} \binom{m}{u} a_2^{m-u}\, \alpha_2^{u}\, V_{u,n} \quad .$$

Thus the mth moment of the distribution on the $(n + 1)$th trial depends on all the moments up to the mth on trial n.

From the above formula,

$$(1.2.8)\quad V_{1,n+1} = \pi_1\, a_1 + \pi_2\, a_2 + (\pi_1\, \alpha_1 + \pi_2\, \alpha_2) V_{1,n}$$

$$(1.2.9)\quad V_{2,n+1} = \pi_1\, a_1^2 + \pi_2\, a_2^2 + 2(\pi_1\, a_1\, \alpha_1 + \pi_2\, a_2\, \alpha_2) V_{1,n}$$

$$+ (\pi_1\, \alpha_1^2 + \pi_2\, \alpha_2^2) V_{2,n}$$

$$= C_0 + C_1\, V_{1,n} + C_2\, V_{2,n} \quad ,$$

where

$$C_o = \pi_1 a_1^2 + \pi_2 a_2^2$$

$$C_1 = 2(\pi_1 a_1 \alpha_1 + \pi_2 a_2 \alpha_2)$$

$$C_2 = \pi_1 \alpha_1^2 + \pi_2 \alpha_2^2 .$$

We may write the equation (1.2.8) in the form

$$V_{1,n+1} = \bar{a} + \bar{\alpha} V_{1,n}$$

where $\bar{a} = \pi_1 a_1 + \pi_2 a_2$

and $\bar{\alpha} = \pi_1 \alpha_1 + \pi_2 \alpha_2$.

This is a simple linear difference equation with general solution

$V_{1,n} = \dfrac{\bar{a}}{1 - \bar{\alpha}} + K \bar{\alpha}^n$, where K is an arbitrary constant.

Let $V_{1,\infty} = \lim\limits_{n \to \infty} V_{1,n} = \dfrac{\bar{a}}{1 - \bar{\alpha}}$.

Then $V_{1,n} = V_{1,\infty} + K \bar{\alpha}^n$,

and since

$$V_{1,o} = V_{1,\infty} + K ,$$

we obtain

$$K = V_{1,o} - V_{1,\infty} .$$

Finally

(1.2.10) $$V_{1,n} = V_{1,\infty} - (V_{1,\infty} - V_{1,o})\bar{\alpha}^n .$$

Substituting the value of $V_{1,n}$ into the formula for $V_{2,n+1}$,

$$(1.2.11) \quad V_{2,n+1} = (C_o + C_1 V_{1,\infty}) - C_1(V_{1,\infty} - V_{1,o})\bar{\alpha}^n + C_2 V_{2,n}$$

$$= C_o' + C_1' \bar{\alpha}^n + C_2 V_{2,n}$$

where $C_o' = C_o + C_1 V_{1,\infty}$

$$C_1' = -C_1(V_{1,\infty} - V_{1,o}) \quad .$$

By definition, $V_{2,o} = V_{1,o}^2 = p_o^2$

where p_o = initial probability of making response A_1 , and so

$$V_{2,1} = C_o' + C_1' + C_2 p_o^2$$

$$V_{2,2} = C_o' + C_1' \bar{\alpha} + C_2(C_o' + C_1' + C_2 p_o^2)$$

$$= C_o'(1 + C_2) + C_1'(C_2 + \bar{\alpha}) + C_2^2 p_o^2 \quad ,$$

and $V_{2,3} = C_o' + C_1' \bar{\alpha}^2 + C_2 V_{2,2}$

$$= C_o' + C_1' \bar{\alpha}^2 + C_2\left[C_o'(1 + C_2) + C_1'(C_2 + \bar{\alpha}) + C_2^2 p_o^2\right]$$

$$= C_o'(1 + C_2 + C_2^2) + C_1'(\bar{\alpha}^2 + C_2 \bar{\alpha} + C_2^2) + C_2^3 p_o^2 \quad .$$

This is of the form

$$(1.2.12) \quad V_{2,n} = C_o' \sum_{u=o}^{n-1} C_2^u + C_1' \sum_{u=o}^{n-1} \bar{\alpha}^{n-1-u} C_2^u + C_2^n p_o^2 \quad .$$

Assume that the above formula holds for the nth case and consider the (n + 1)th case:

$$(1.2.13) \quad V_{2,n+1} = C_o' + C_1' \bar{\alpha}^n + C_2 V_{2,n}$$

$$= C_0' + C_1' \bar{\alpha}^n + C_2 \left(C_0' \sum_{u=0}^{n-1} C_2^u + C_1' \sum_{u=0}^{n-1} \bar{\alpha}^{n-1-u} C_2^u + C_2^n p_0^2 \right.$$

$$= C_0' \left[1 + C_2 \sum_{u=0}^{n-1} C_2^u \right] + C_1' \left[\bar{\alpha}^n + \sum_{u=0}^{n-1} \bar{\alpha}^{n-1-u} C_2^{u+1} \right] + C_2 C_2^n p_0^2$$

$$= C_0' \sum_{u=0}^{n} C_2^u + C_1' \sum_{u=0}^{n} \bar{\alpha}^{n-u} C_2^u + C_2^{n+1} p_0^2 .$$

This is of the same form as in the nth case and so in general

$$(1.2.14) \quad V_{2,n} = C_0' \sum_{u=0}^{n-1} C_2^u + C_1' \sum_{u=0}^{n-1} \bar{\alpha}^{n-1-u} C_2^u + C_2^n p_0^2$$

$$= \frac{C_0'(1 - C_2^n)}{1 - C_2} + \frac{C_1'(\bar{\alpha}^n - C_2^n)}{\bar{\alpha} - C_2} + C_2^n p_0^2$$

$$= (C_0 + C_1 V_{1,\infty}) \frac{1 - C_2^n}{1 - C_2} - C_1 (V_{1,\infty} - p_0) \frac{\bar{\alpha}^n - C_2^n}{\bar{\alpha} - C_2} + C_2^n p_0^2 .$$

Thus for $|a_i| < 1$ and $|\alpha_i| < 1$ $0 \le \pi_i \le 1$,

$$(1.2.15) \quad \lim_{n \to \infty} V_{2,n} = V_{2,\infty} = \frac{C_0 + C_1 V_{1,\infty}}{1 - C_2}$$

and the variance

$$(1.2.16) \quad \sigma^2 = V_{2,\infty} - V_{1,\infty}^2 = \frac{C_0 + C_1 V_{1,\infty}}{1 - C_2} - V_{1,\infty}^2 .$$

An important theorem on the p-value distribution is given by Karlin [2]. It is stated here without proof. For two response classes and experimenter-controlled events, an asymptotic distribution exists, that is, the p-value distributions on trials n and $n + 1$ may be made as close to one another as one wishes by making n sufficiently large. Furthermore, this asymptotic distribution is independent of the initial distribution of p-values.

Section Three

Subject-Controlled Events

If the occurrence of a response is taken as an event, then events of this type are called subject-controlled events. For example, for two alternative responses A_1 and A_2, the experiment is subject-controlled if event E_1 occurs whenever the response A_1 occurs and event E_2 occurs whenever the response A_2 occurs. In the Brunswik T-Maze experiment, let it be so arranged that if the rat turns right, reward is always found and if the rat turns left, reward is never found. Then this simple T-Maze experiment is a subject-controlled experiment if the response A_1 of the rat to turn right is taken as event E_1 and the response A_2 of the rat to turn left is taken as event E_2. In the two experimenter-controlled event case, the response is determined by the animal, while the outcome (i.e. reward or non-reward) is determined by the experimenter.

However, in the case of the subject-controlled events, the outcome is completely determined by the response. So in the case of two subject-controlled events, event E_1 occurs whenever response A_1 occurs and event E_2 occurs whenever response A_2 occurs.

Under the above definition of subject controlled-events, the probability of the occurrence of event E_1 on trial n is not a constant as in the case of experimenter-controlled events, but is equal to the probability p_n of response A_1. If we let the possible values of the response probability p be the states of the system, the process is a Markov process since the conditional probabilities of all other states, given state p, are independent of the way in which the state p was achieved i.e. since if p_n is the probability of A_1 on trial n the possible values of p_{n+1} depend only on p_n and not on $p_o, \ldots, p_{n-1}$.

Since the probability of applying Q_1 to p_n is p_n and the probability of applying Q_2 to p_n is $1 - p_n$, a branching process similar to the case of two experimenter-controlled events is obtained and is shown by figure 1.6.

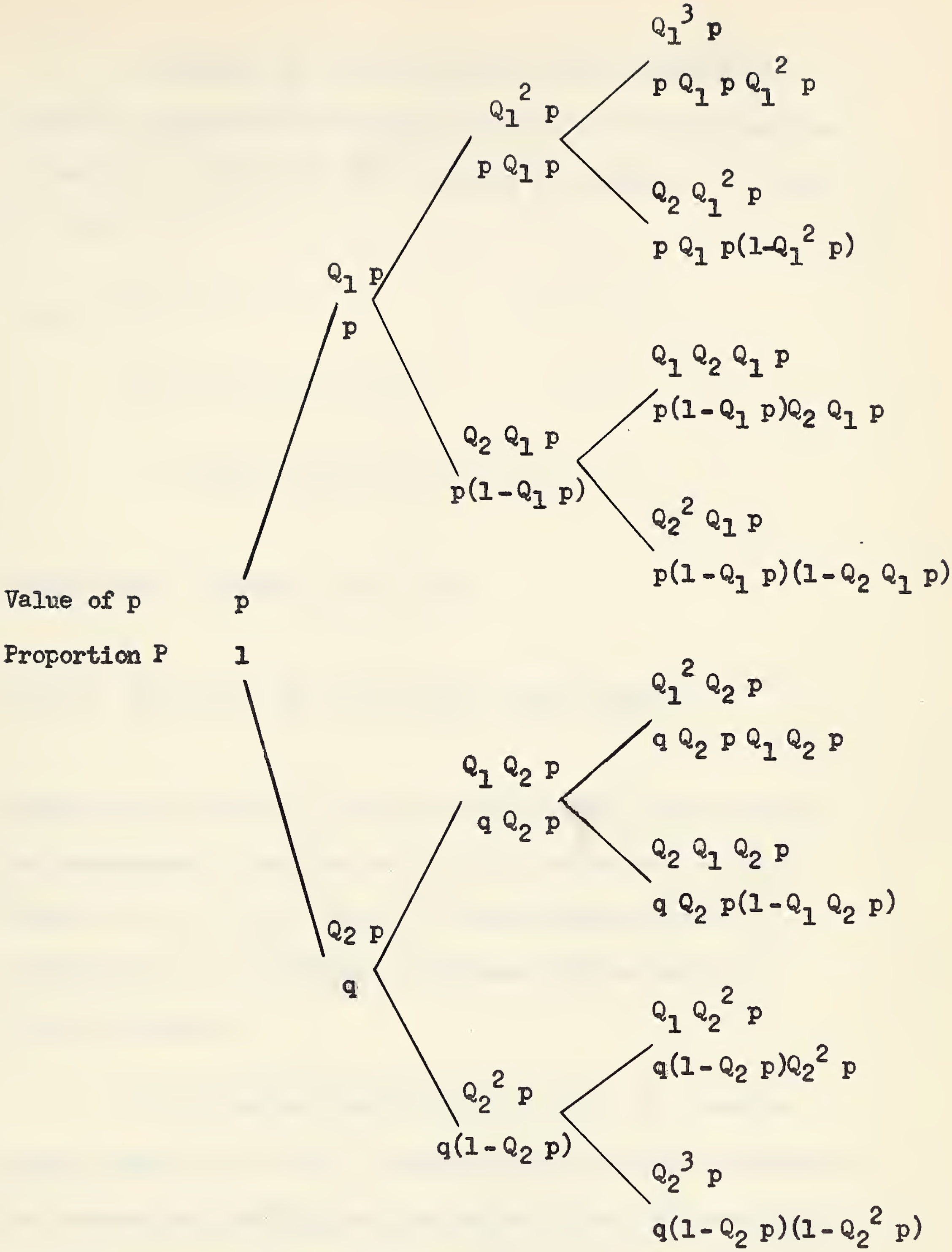

Fig. 1.6 This diagram shows the successive splits that a large number of organisms would go through on successive applications of the operators for the case of subject-controlled events. Both the p value of the group and the proportional size of the group are given at each stage of the operation. The probability of applying Q_1 is equal to the p value of the group. (Reproduced from (1) page 79)

Defining $\bar{p}_n$ as in the case of experimenter-controlled events as the average probability of the occurrence of event E_1 on the $(n+1)$th trial, the values of $\bar{p}_0$ and $\bar{p}_1$ are

$$(1.3.1) \quad \bar{p}_0 = p \ ,$$

$$\bar{p}_1 = p\,Q_1\,p + q\,Q_2\,p$$

$$= a_1 p + \alpha_1 p^2 + a_2 q + \alpha_2 p q$$

respectively. Putting $q = 1 - p$,

$$(1.3.2) \quad \bar{p}_1 = a_2 + (a_1 - a_2 + \alpha_2)p + (\alpha_1 - \alpha_2)p^2 \ ,$$

which is not linear in p and so is not in the form in which the mathematics of section one of this chapter can be used. However, if $\alpha_1 = \alpha_2$, the p^2 term disappears and the expression in p is linear. This case will be left to a later discussion.

It is obvious from figure 1.6, that $\bar{p}_n$ involves higher powers of p as n increases and this fact complicates the mathematics involved and the methods described in section one of this chapter cannot be used.

Moments for Two Subject-Controlled Events

Let $p_{\nu n}$ be the p-value for the νth group of

organisms after n applications of the operators, and let $P_{\nu n}$ be the proportion of the νth group in the whole population. Upon a further application of the operators, the following is obtained:

New p-Value	New Proportion
$Q_1 p_{\nu n} = a_1 + \alpha_1 p_{\nu n}$	$p_{\nu n} P_{\nu n}$
$Q_2 p_{\nu n} = a_2 + \alpha_2 p_{\nu n}$	$(1 - p_{\nu n})P_{\nu n}$.

Thus by the definition of moments,

$$(1.3.3)\quad V_{m,n+1} = \sum_{\nu=1}^{2^n} (Q_1 p_{\nu n})^m p_{\nu n} P_{\nu n} + \sum_{\nu=1}^{2^n} (Q_2 p_{\nu n})^m (1 - p_{\nu n}) P_{\nu n}$$

$$= \sum_{\nu=1}^{2^n} (a_1 + \alpha_1 p_{\nu n})^m p_{\nu n} P_{\nu n} + \sum_{\nu=1}^{2^n} (a_2 + \alpha_2 p_{\nu n})^m (1 - p_{\nu n}) P_{\nu n}$$

$$= \sum_{\nu=1}^{2^n} \sum_{u=0}^{m} \binom{m}{u} a_1^{m-u} \alpha_1^{u} p_{\nu n}^{u+1} P_{\nu n} - \sum_{\nu=1}^{2^n} \sum_{u=0}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{u} p_{\nu n}^{u+1} P_{\nu n}$$

$$+ \sum_{\nu=1}^{2^n} \sum_{u=0}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{u} p_{\nu n}^{u} P_{\nu n}$$

$$= \sum_{u=0}^{m} \binom{m}{u} a_1^{m-u} \alpha_1^{u} V_{u+1,n} - \sum_{u=0}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{u} V_{u+1,n}$$

$$+ \sum_{u=0}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{u} V_{u,n}$$

$$= \sum_{u=1}^{m+1} \binom{m}{n-1} (a_1^{m-u+1} \; \alpha_1^{u-1} - a_2^{m-u+1} \; \alpha_2^{u-1}) V_{u,n}$$

$$+ \sum_{u=0}^{m} \binom{m}{u} a_2^{m-u} \; \alpha_2^{u} \; V_{u,n}$$

$$= a_2^{m} \; V_{o,n} + \sum_{u=1}^{m} \left[\binom{m}{u-1} (a_1^{m-u+1} \; \alpha_1^{u-1} - a_2^{m-u+1} \; \alpha_2^{u-1}) \right] V_{u,n}$$

$$+ \sum_{u=1}^{m} \binom{m}{u} a_2^{m-u} \; \alpha_2^{u} \; V_{u,n} + (\alpha_1^{m} - \alpha_2^{m}) V_{m+1,n}$$

$$= \sum_{u=o}^{m+1} C_{m,u} \; V_{u,n} \qquad ,$$

where

$$C_{m,u} = \begin{cases} a_2^{m} & (u = 0) \\ \binom{m}{u-1} (a_1^{m-u+1} \; \alpha_1^{u-1} - a_2^{m-u+1} \; \alpha_2^{u-1}) + \binom{m}{u} a_2^{m-u} \; \alpha_2^{u} & \\ & (u = 1, 2, \ldots m) \\ \alpha_1^{m} - \alpha_2^{m} & (u = m + 1) \end{cases}$$

.

Thus for the mean on trial $n + 1$

(1.3.4) $$V_{1,n+1} = C_{10} + C_{11} V_{1,n} + C_{12} V_{2,n}$$

where

$$C_{10} = a_2$$

$$C_{11} = a_1 - a_2 + \alpha_2$$

$$C_{12} = \alpha_1 - \alpha_2$$

and for the second raw moment,

(1.3.5) $$V_{2,n+1} = C_{20} + C_{21} V_{1,n} + C_{22} V_{2,n} + C_{23} V_{3,n}$$

where

$$C_{20} = a_2^2$$

$$C_{21} = a_1^2 - a_2^2 + 2a_2 \alpha_2$$

$$C_{22} = 2(a_1 \alpha_1 - a_2 \alpha_2) + \alpha_2^2$$

$$C_{23} = \alpha_1^2 - \alpha_2^2 \quad .$$

It can be seen from the recurrence relation (1.3.3)

$V_{m,n+1} = \sum_{u=0}^{m+1} C_{m,u} V_{u,n}$ that the evaluation of $V_{m,n+1}$ requires all the moments through the $(m + 1)th$ on

the nth trial and this makes the higher moments awkward to evaluate.

Equal α-Condition

Consider the case $\alpha_1 = \alpha_2 = \alpha$. Then the formula (1.3.4)

$$V_{1,n+1} = C_{10} + C_{11} V_{1,n} + C_{12} V_{2,n} \quad \text{with}$$

$$C_{10} = a_2 \ , \quad C_{11} = a_1 - a_2 + \alpha_2 \ , \quad \text{and} \quad C_{12} = \alpha_1 - \alpha_2$$

becomes

$$(1.3.5) \quad V_{1,n+1} = a_2 + (a_1 - a_2 + \alpha_2)V_{1,n} + (\alpha_1 - \alpha_2)V_{2,n}$$

$$= a_2 + (a_1 - a_2 + \alpha)V_{1,n} \ ,$$

which is of the same form as the slope intercept form. This suggests the introduction of the operator $\bar{Q}$ such that

$$(1.3.6) \quad V_{1,n+1} = \bar{Q}\, V_{1,n} = a_2 + (a_1 - a_2 + \alpha)V_{1,n}$$

$$= a_2 + \beta\, V_{1,n} \ ,$$

where $\beta = a_1 - a_2 + \alpha$.

Thus, by comparing with the slope intercept form,

$$a_2 = (1 - \beta)V_{1,\infty} \qquad \text{for} \quad \beta \neq 1 \ ,$$

and so the asymptotic formula for the mean is

(1.3.7) $$V_{1,\infty} = \frac{a_2}{1 - \beta} = \frac{a_2}{1 - a_1 + a_2 - \alpha} = \frac{\lambda_2}{1 - \lambda_1 + \lambda_2} \quad .$$

Furthermore, the explicit formula for the means can be obtained as

(1.3.8) $$V_{1,n} = \overline{Q}^n \, V_{1,o} = \beta^n \, V_{1,o} + (1 - \beta^n)V_{1,\infty} \quad .$$

An interesting property of the mean of the two subject-controlled events distribution is that it can be expressed independently of the value of α .

Putting $V_{1,n+1} = a_2 + (a_1 - a_2 + \alpha)V_{1,n}$ in the gain-loss form of Q_1 by introducing the parameter $b_1 = 1 - a_1 - \alpha$ as defined in section one,

(1.3.9) $$V_{1,n+1} = a_2 + (1 - a_2 - b_1)V_{1,n} \quad .$$

Thus for the case where the α's are equal, the means are independent of the value of α .

The recurrence formula for the moments is

$$V_{m,n+1} = \sum_{u=o}^{m} \binom{m}{u} a_1^{m-u} \alpha_1^{\ u} V_{u+1,n} - \sum_{u=o}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{\ u} V_{u+1,n}$$

$$+ \sum_{u=o}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{\ u} V_{u,n} .$$

Putting $\alpha_1 = \alpha_2 = \alpha$ this becomes

(1.3.10) $$V_{m,n+1} = \sum_{u=o}^{m} \binom{m}{u} \left[a_1^{m-u} \alpha^u V_{u+1,n} + a_2^{m-u} \alpha^u (V_{u,n} - V_{u+1,n}) \right]$$

$$= \sum_{u=o}^{m-1} \binom{m}{u} \left[a_1^{m-u} \alpha^{u} V_{u+1,n} + a_2^{m-u} \alpha^u (V_{u,n} - V_{u+1,n}) \right]$$

$$+ \alpha^m V_{m,n} .$$

Thus for the second moment,

(1.3.11) $$V_{2,n+1} = a_1^{\ 2} V_{1,n} + a_2^{\ 2}(1 - V_{1,n})$$

$$+ 2\left[a_1 \alpha V_{2,n} + a_2 \alpha (V_{1,n} - V_{2,n}) \right] + \alpha^2 V_{2,n}$$

$$= a_2^2 + (a_1^{\ 2} - a_2^{\ 2} + 2a_2 \alpha) V_{1,n} + \alpha(2a_1 - 2a_2 + \alpha) V_{2,n}$$

$$= a_2^{\ 2} + B_1 V_{1,n} + B_2 V_{2,n} ,$$

where $$B_1 = a_1^2 - a_2^2 + 2a_2\,\alpha$$

$$B_2 = \alpha(2a_1 - 2a_2 + \alpha) \quad .$$

Since $$V_{1,n} = V_{1,\infty} - (V_{1,\infty} - V_{1,0})\beta^n \quad ,$$

$$V_{2,n+1} = (a_2^2 + B_1 V_{1,\infty}) - B_1(V_{1,\infty} - V_{1,0})\beta^n + B_2 V_{2,n} \tag{1.3.12}$$

which is a difference equation of the same form as (1.2.13) in the case of two experimenter-controlled events, and so, applying the same method of solution,

$$V_{2,n} = (a_2^2 + B_1 V_{1,\infty}) \frac{1 - B_2^n}{1 - B_2} - B_1(V_{1,\infty} - V_{1,0}) \frac{\beta^n - B_2^n}{\beta - B_2} + B_2^n V_{2,0} \quad , \tag{1.3.13}$$

where $V_{1,0}$ and $V_{2,0}$ are the initial mean and second raw moments respectively. If $-1 < B_2 < 1$, $-1 < \beta < 1$, and $\beta \neq B_2$, in the limit as $n \to \infty$, the asymptotic formula for the second moment is

$$V_{2,\infty} = \frac{a_2^2 + B_1 V_{1,\infty}}{1 - B_2} \quad , \tag{1.3.14}$$

where $V_{1,\infty}$ is the asymptotic formula for the mean. Thus the variance of the two subject-controlled events distribution in the equal α case is

$$(1.3.15) \qquad \sigma^2 = V_{2,\infty} - V_{1,\infty}^2$$

$$= \frac{a_2^{\,2} + B_1 V_{1,\infty}}{1 - B_2} - V_{1,\infty}^2 \ .$$

Asymptotic Distribution Theorem.

Harris has shown in [3] that the same asymptotic distribution theorem for two response classes and experimenter-controlled events stated at the end of section two of this chapter holds for two subject-controlled events if the α_i's are non-negative and the absolute value of the difference between the two limit points is less than unity. When the limits are zero and unity, these special cases are discussed in [1] chapter seven. For the case $\lambda_1 = 1$, and $\lambda_2 = 0$, it is shown that the asymptotic distribution depends on the initial probability p_0 except when $\alpha_1 = \alpha_2 = 1$. For the case $\lambda_1 = 0$ and $\lambda_2 = 1$, it is shown that the asymptotic distribution is independent of the initial probability.

CHAPTER II

MOMENT GENERATING FUNCTIONAL EQUATIONS ASSOCIATED WITH TWO EXPERIMENTER-CONTROLLED EVENTS AND OF TWO SUBJECT-CONTROLLED EVENTS

In this chapter, the moment generating functional equations of the distributions of p-values for two experimenter-controlled events and two subject-controlled events will be derived. In section one we review the techniques used by R. R. Bush and F. Mosteller for studying the asymptotic distributions of p-values. Section two of this chapter deals with the experimenter-controlled events, and section three deals with the subject-controlled events.

Section One

An approximation to the asymptotic distribution can be obtained by the "Stat-Rat" or "Monte Carlo" method when the constants of the operators are given. A stat-rat, in fact, is a "hypothetical animal" which obeys the given mathematical model. The method makes use of random numbers and we illustrate it by the following example.

Let the experiment be subject-controlled with the mathematical model $p_0 = 0.2$, $a_1 = 0.3$, $\alpha_1 = 0.6$, $a_2 = 0.01$, $\alpha_2 = 0.9$ i.e. $Q_1 p = 0.6p + 0.3$ and $Q_2 p = 0.9p + 0.01$.

Table 2.1 gives the mechanism of the method. Column one gives the trial number; the second column gives the p-values for each trial and the third column is a set of 2-digit numbers from a random numbers table and the fourth column gives the operator to be applied on each trial. The random numbers of column three are of the form 00, 01, ..., 99. The number 00 stands for all decimal numbers from 0.00000... to 0.00999... respectively. Operator Q_1 is applied whenever the random numbers stand for values less than the corresponding p-value on the nth trial. Operator Q_2 is applied otherwise. For example, if the probability of applying Q_1 is 0.344 on a certain trial, then if the random number taken is less than 34 i.e. 00, 01, ..., 33, operator Q_1 is applied. If the random number is greater than 34, i.e. 35, 36, ... 99, operator Q_2 is then applied. If the random number 34 is obtained which is equal to the first 2 digits of the probability 0.344, the ambiguity is solved by adding additional digits to the end of the number until a decision is reached. The above procedure is clarified if we follow a few steps in table 2.1. The initial p-value was 0.2 and the first random number taken was 84 which is greater than 20 and so Q_2 was applied to the initial p (0.2) to give the new p value 0.1900. The next random number taken was 29 which was greater than 19 and so Q_2 was applied again. This procedure was carried on until on the 6th trial, the random number 05 which is less than 15 was chosen. So Q_1 was applied to 0.1531 giving a new p-value of 0.3919.

TABLE 2.1

Illustration of a computation sheet for 25 trials of one stat-rat. The operations are $Q_1p = 0.3 + 0.6p$ and $Q_2p = 0.01 + 0.9p$ and $p_o = 0.2$. (Reproduced from (1) page 130)

Trial Number	p-value	Random Number	Operator
0	0.2000	84	Q_2
1	0.1900	29	Q_2
2	0.1810	35	Q_2
3	0.1729	69	Q_2
4	0.1656	53	Q_2
5	0.1590	37	Q_2
6	0.1531	05	Q_1
7	0.3919	50	Q_2
8	0.3627	57	Q_2
9	0.3364	60	Q_2
10	0.3128	55	Q_2
11	0.2915	58	Q_2
12	0.2724	79	Q_2
13	0.2552	50	Q_2
14	0.2397	56	Q_2
15	0.2257	01	Q_1
16	0.4354	51	Q_2
17	0.4019	65	Q_2
18	0.3717	92	Q_1
19	0.3445	32	Q_1
20	0.5067	21	Q_2
21	0.6040	66	Q_1
22	0.5536	35	Q_1
23	0.6322	18	Q_1
24	0.6793	65	Q_1
25	0.7076		

Although table 2.1 gives the run of a stat-rat of 25 trials, the number of trials can be carried as far as one wishes. When the number of trials is sufficiently large, the final p-value can be taken as an approximation to the asymptotic p-value. The initial p-value can be taken arbitrarily since the asymptotic p-values are, except in certain special cases stated at the end of section three, chapter one, independent of the initial p-value. Thus by taking a large number of stat-rats, an approximation of the asymptotic distribution can be obtained. Naturally, we would like to approximate the asymptotic distribution with a specified accuracy. Suppose that we want to have a class interval γ in computing the approximate distribution. As p_o is arbitrary, the question arises how many trials are necessary in order that p_n and p_n' arising from p_o and p_o' lie within the same class interval. That is, we want to find an n such that

$$(2.1.1) \qquad p_n - p_n' \leq \gamma \quad .$$

For the experimenter-controlled events,

$$Q_i\, p_o - Q_i\, p_o' = \alpha_i (p_o - p_o') \quad .$$

Let β be the largest α_i $(i = 1, \ldots t)$, and $p_o > p_o'$. Then

$$Q_i\, p_o - Q_i\, p_o' \leq \beta(p_o - p_o') \quad ,$$

i.e. $\quad p_1 - p_1' \leq \beta(p_o - p_o')$

and $\quad p_2 - p_2' \leq \beta^2(p_o - p_o')$.

Thus by induction, in general

$$(2.1.3) \quad p_n - p_n' \leq \beta^n(p_o - p_o') \quad .$$

It can be shown (see "Trapping Theorem" [1] page 99) that the asymptotic p-values lie between $\lambda_{min.}$ and $\lambda_{max.}$ Thus choosing p_o and p_o' to lie between these limits it follows that $p_o - p_o'$ is less than or equal to $(\lambda_{max.} - \lambda_{min.})$, and so we would require n to satisfy

$$(2.1.4) \quad \beta^n(\lambda_{max.} - \lambda_{min.}) \leq \gamma \quad .$$

Thus if $\alpha_1 = 0.6$, $\alpha_2 = 0.9$, $\lambda_1 = 0.75$, $\lambda_2 = 0.1$, then $\beta = 0.9$, $\lambda_{max.} = 0.75$ and $\lambda_{min.}$ is 0.1. If the class interval $\gamma = 0.01$ the inequality is

$$(2.1.5) \quad (0.9)^n\,(0.75 - 0.10) \leq 0.01$$

$$\text{or} \qquad (0.9)^n \leq 0.0154 \quad .$$

The smallest n which satisfies this inequality is $n = 40$. Thus 40 stat-rat trials would make certain that even the extreme values of p_o lead to final p-values in the same of adjacent classes in the desired distribution. However, in ordinary cases, a smaller number is required as it is rare to use $\lambda_{max.}$ and $\lambda_{min.}$ values as p_o .

TABLE 2.2

Frequencies of p values which occurred in a 1000 trial stat-rat with $Q_1p = 0.3 + 0.6p$, $Q_2p = 0.01 + 0.9p$, $\pi_1 = \pi_2 = 0.5$. All p values were rounded to two decimals. (Reproduced from (1) page 134)

p	f	p	f	p	f
≤ 0.26	0	0.43	6	0.60	33
0.27	1	0.44	5	0.61	53
0.28	0	0.45	7	0.62	58
0.29	1	0.46	7	0.63	23
0.30	0	0.47	13	0.64	58
0.31	1	0.48	10	0.65	44
0.32	1	0.49	11	0.66	36
0.33	1	0.50	8	0.67	65
0.34	1	0.51	28	0.68	57
0.35	0	0.52	17	0.69	28
0.36	1	0.53	20	0.70	55
0.37	2	0.54	28	0.71	44
0.38	2	0.55	18	0.72	30
0.39	6	0.56	32	0.73	32
0.40	3	0.57	36	0.74	46
0.41	3	0.58	22	0.75	0
0.42	9	0.59	38		1000

For two subject-controlled events, it is difficult to obtain a good estimate of this kind.

The above method, however, is very wasteful because only the final p-value of a stat-rat is used for the approximation of the asymptotic distribution. It can be shown (see "Ergodic Theorem" [1] page 99) that a single stat-rat can generate the asymptotic distribution provided the sequence becomes infinitely long. Thus as an approximation, if the total number of trials is large compared with the number necessary for a single final value, the stat-rat can be used to approximate the asymptotic distribution. For example, in the previous numerical case, instead of 40 trials, a single stat-rat of say 1000 trials would be quite adequate and the method of the approximation is illustrated by table 2.2.

By the stat-rat method, R. R. Bush and F. Mosteller in [1] obtain a number of approximate asymptotic distributions, but the only exact asymptotic distribution which they have obtained is the case of two subject-controlled events with $\alpha_1 = \alpha_2 = \alpha$ and $\lambda_1 = 1$, $\lambda_2 = 0$. It is a simple binomial distribution with all density concentrated at $p = 0$ and at $p = 1$.

Section Two

We define a moment generating functional equation as follows. Let x be a random variable and let $M_x(t)$ be the moment generating function of the probability distribution of x. Then a moment generating functional equation is a functional equation satisfied by $M_x(t)$. For example, if x is a binomial random variable taking values -1 and $+1$ with probabilities $\frac{1}{2}$ and $\frac{1}{2}$ respectively, then

$$(2.2.1) \qquad M_x(t) = \cosh(t) \quad , \qquad \text{and since}$$

$$(2.2.2) \qquad \cosh 2t = 2\cosh^2 t - 1 \quad ,$$

a moment generating functional equation satisfied by $M_x(t)$ is

$$(2.2.3) \qquad M_x(2t) = 2 \cosh t \, M_x(t) - 1 \quad .$$

In chapter one, it was shown that on the nth trial of a two experimenter-controlled events experiment, there are 2^n possible probabilities $p_{\nu n}$ (called p-values) ($\nu = 1, \ldots 2^n$) of making response A_1 each with proportion $P_{\nu n}$. Thus the moment generating function of the distribution of p-values on the nth trial is

$$(2.2.4) \qquad M_n(t) = \sum_{\nu=1}^{2^n} e^{tp_{\nu n}} P_{\nu n} \quad .$$

Since each $p_{\nu n}$ on trial "n" generates two new values $a_1 + \alpha_1 p_{\nu n}$ with probability $\pi_1 p_{\nu n}$ and $a_2 + \alpha_2 p_{\nu n}$ with probability $\pi_2 p_{\nu n}$, on trial "n + 1", it follows that

$$(2.2.5)\quad M_{n+1}(t) = \sum_{\nu=1}^{2^n} \left(e^{t(a_1+\alpha_1 p_{\nu n})} \pi_1 p_{\nu n} + e^{t(a_2+\alpha_2 p_{\nu n})} \pi_2 p_{\nu n}\right)$$

$$= \pi_1 e^{ta_1} \sum_{\nu=1}^{2^n} e^{(t\alpha_1)p_{\nu n}} p_{\nu n} + \pi_2 e^{ta_2} \sum_{\nu=1}^{2^n} e^{(t\alpha_2)p_{\nu n}} p_{\nu n}$$

$$= \pi_1 e^{ta_1} M_n(\alpha_1 t) + \pi_2 e^{ta_2} M_n(\alpha_2 t) \quad ,$$

and so as n approaches infinity, we obtain the asymptotic functional equation

$$(2.2.6)\quad M(t) = \pi_1 e^{ta_1} M(\alpha_1 t) + \pi_2 e^{ta_2} M(\alpha_2 t)$$

where $M(t) \equiv M_\infty(t)$, $M(\alpha_1 t) \equiv M_\infty(\alpha_1 t)$ and $M(\alpha_2 t) \equiv M_\infty(\alpha_2 t)$.

In the case $\alpha_1 = \alpha_2 = \alpha$

$$(2.2.7)\quad M_{n+1}(t) = M_n(\alpha t)\,[\pi_1 e^{ta_1} + \pi_2 e^{ta_2}] \quad , \quad \text{and}$$

$$(2.2.8)\quad M(t) = M(\alpha t)\,[\pi_1 e^{ta_1} + \pi_2 e^{ta_2}] \quad .$$

Section Three

From chapter one, on the "n" th trial of the two subject-controlled events experiment, there are 2^n possible p-values $p_{\nu n}$ (where $\nu = 1, \ldots 2^n$) of making response A_1 each with proportion $P_{\nu n}$ respectively. Thus the moment generating function of the distribution of p-values on the "n" th trial is

$$(2.3.1) \qquad M_n(t) = \sum_{\nu=1}^{2^n} e^{tp_{\nu n}} P_{\nu n} \quad .$$

Since each $p_{\nu n}$ on trial "n" generates two new values

$$a_1 + \alpha_1 p_{\nu n} \quad \text{with probability} \quad p_{\nu n} P_{\nu n} \quad ,$$

and $$a_2 + \alpha_2 p_{\nu n} \quad \text{with probability} \quad (1 - p_{\nu n})P_{\nu n} \quad ,$$

on trial "n + 1" it follows that

$$(2.3.2) \qquad M_{n+1}(t) = \sum_{\nu=1}^{2^n} e^{t(a_1+\alpha_1 p_{\nu n})} p_{\nu n} P_{\nu n} + \sum_{\nu=1}^{2^n} e^{t(a_2+\alpha_2 p_{\nu n})}(1 - p_{\nu n})P_{\nu n}$$

$$= e^{ta_1} \sum_{\nu=1}^{2^n} e^{t\alpha_1 p_{\nu n}} p_{\nu n} P_{\nu n} + e^{ta_2} \sum_{\nu=1}^{2^n} e^{t\alpha_2 p_{\nu n}} P_{\nu n}$$

$$-e^{ta_2} \sum_{\nu=1}^{2^n} e^{t\alpha_2 p_{\nu n}} p_{\nu n} P_{\nu n}$$

$$= e^{ta_2} \sum_{\nu=1}^{2^n} e^{t\alpha_2 p_{\nu n}} P_{\nu n} + \frac{e^{ta_1}}{\alpha_1} \frac{d}{dt} \sum_{\nu=1}^{2^n} e^{t\alpha_1 p_{\nu n}} P_{\nu n}$$

$$- \frac{e^{ta_2}}{\alpha_2} \frac{d}{dt} \sum_{\nu=1}^{2^n} e^{t\alpha_2 p_{\nu n}} P_{\nu n}$$

$$= e^{ta_2} M_n(\alpha_2 t) + \frac{e^{ta_1}}{\alpha_1} \frac{d}{dt} M_n(\alpha_1 t) - \frac{e^{ta_2}}{\alpha_2} \frac{d}{dt} M_n(\alpha_2 t) ,$$

where $\alpha_1 , \alpha_2 \neq 0$. For the equal α case,

$$(2.3.3) \quad M_{n+1}(t) = e^{ta_2} M_n(\alpha t) + \frac{e^{ta_1} - e^{ta_2}}{\alpha} \frac{d}{dt} M_n(\alpha t) .$$

Letting n approach infinity, the asymptotic moment-generating functional equation is

$$(2.3.4) \lim_{n \to \infty} M_{n+1}(t) = M_\infty(t) = e^{ta_2} M_\infty(\alpha_2 t) + \frac{e^{ta_1}}{\alpha_1} \frac{d}{dt} M_\infty(\alpha_1 t) ,$$

$$- \frac{e^{ta_2}}{\alpha_2} \frac{d}{dt} M_\infty(\alpha_2 t)$$

$$\text{or} \quad M(t) = e^{ta_2} M(\alpha_2 t) + \frac{e^{ta_1}}{\alpha_1} \frac{d}{dt} M(\alpha_1 t) - \frac{e^{ta_2}}{\alpha_2} \frac{d}{dt} M(\alpha_2 t) ,$$

where $M(t) \equiv M_\infty(t)$, $M(\alpha_1 t) \equiv M_\infty(\alpha_1 t)$ and $M(\alpha_2 t) \equiv M_\infty(\alpha_2 t)$.

For the equal α case, the asymptotic moment generating functional equation is

$$M(t) = e^{ta_1} M(\alpha t) + \frac{e^{ta_1} - e^{ta_2}}{\alpha} \frac{d}{dt} M(\alpha t) \quad , \tag{2.3.5}$$

We define the characteristic function of an asymptotic distribution to be

$$\phi(t) = M(it) , \qquad \text{where} \quad i^2 = -1 \quad . \tag{2.3.6}$$

Then from (2.3.4) the characteristic function of the asymptotic distribution of the two subject - controlled events model satisties the functional equation

$$\phi(t) = e^{ita_1} \phi(\alpha_1 t) - \frac{ie^{ita_1}}{\alpha_1} \frac{d}{dt} \phi(\alpha_1 t) + \frac{ie^{ita_2}}{\alpha_2} \frac{d}{dt} \phi(\alpha_2 t) \tag{2.3.7}$$

In the equal α case, (2.3.7) becomes

$$\phi(t) = e^{ita_2} \phi(\alpha t) - \frac{i(e^{ita_1} - e^{ita_2})}{\alpha} \frac{d}{dt} \phi(\alpha t) \quad , \tag{2.3.8}$$

and if $\alpha = \frac{1}{2}$ we have

$$\phi(t) = e^{ita_2} \phi(t/2) - 2i \left(e^{ita_1} - e^{ita_2} \right) \frac{d}{dt} \phi(t/2) \quad . \tag{2.3.9}$$

CHAPTER III

SOLUTION OF SOME SPECIAL CASES OF THE ASYMPTOTIC MOMENT GENERATING FUNCTIONAL EQUATIONS

In this chapter, some special cases of the moment generating functional equations under the equal α condition derived in chapter two will be solved and their properties will be discussed. The first section deals with two experimenter-controlled events and the second section deals with two subject-controlled events.

Section One

From chapter two, equation (2.2.7), the moment generating function of the two experimenter-controlled events model in the equal α case satisfies

$$M_{n+1}(t) = M_n(\alpha t)\left[\pi_1 e^{ta_1} + \pi_2 e^{ta_2}\right] .$$

Since $$M_1(t) = \pi_1 e^{t(a_1+\alpha p)} + \pi_2 e^{t(a_2+\alpha p)}$$

$$= e^{\alpha pt}(\pi_1 e^{ta_1} + \pi_2 e^{ta_2}) ,$$

where p = initial probability of making response A_1 ,

$$M_2(t) = M_1(\alpha t)(\pi_1 e^{ta_1} + \pi_2 e^{ta_2})$$

$$= e^{\alpha^2 pt} \prod_{j=0}^{1} (\pi_1 e^{t\alpha^j a_1} + \pi_2 e^{t\alpha^j a_2})$$

$$M_3(t) = M_2(\alpha t)(\pi_1 e^{ta_1} + \pi_2 e^{ta_2})$$

$$= e^{\alpha^3 pt} \prod_{j=0}^{2} (\pi_1 e^{ta_1\alpha^j} + \pi_2 e^{ta_2\alpha^j}) ,$$

the asymptotic moment generating function is equal to

$$(3.1.1) \qquad M(t) = \lim_{n \to \infty} e^{\alpha^n pt} \prod_{j=0}^{n-1} (\pi_1 e^{ta_1\alpha^j} + \pi_2 e^{ta_2\alpha^j})$$

$$= \prod_{j=0}^{\infty} (\pi_1 e^{ta_1\alpha^j} + \pi_2 e^{ta_2\alpha^j}) \quad \text{for} \quad 0 \leq \alpha < 1 .$$

To see if $M(t)$ for $0 \leq \alpha < 1$ is convergent or not, assume

(a) $0 \leq \pi_1 , \pi_2 \leq 1$ and $\pi_1 + \pi_2 = 1$,

(b) $0 \leq a_1 , a_2 \leq 1$ and $a_1 \geq a_2$,

(c) $0 \leq \alpha < 1$,

and (d) t is real .

Then

$$(3.1.2) \qquad \left| \log_e M(t) \right| = \left| \sum_{j=0}^{\infty} \log_e (\pi_1 e^{ta_1\alpha^j} + \pi_2 e^{ta_2\alpha^j}) \right| .$$

Let

$$-\infty < t < 0 \quad , \quad \text{i.e.} \quad t = - |t| \quad .$$

Then $\left| \log_e M(t) \right|$ satisfies

$$(3.1.3) \qquad \left| \log_e M(t) \right| \leq \sum_{j=0}^{\infty} \left| \log_e (\pi_1 e^{-|t| a_1 \alpha^j} + \pi_2 e^{-|t| a_2 \alpha^j}) \right|$$

$$\leq \sum_{j=0}^{\infty} \left| \log_e \left(\frac{\pi_1 e^{|t| a_2 \alpha^j} + \pi_2 e^{|t| a_1 \alpha^j}}{e^{|t| (a_1 + a_2) \alpha^j}} \right) \right|$$

$$\leq \sum_{j=0}^{\infty} \log_e (\pi_1 e^{|t| a_2 \alpha^j} + \pi_2 e^{|t| a_1 \alpha^j})$$

$$+ \sum_{j=0}^{\infty} \log_e e^{|t| (a_1 + a_2) \alpha^j}$$

$$\leq \sum_{j=0}^{\infty} \left| \log_e (\pi_1 e^{|t| a_1 \alpha^j} + \pi_2 e^{|t| a_1 \alpha^j}) \right|$$

$$+ \sum_{j} |t| (a_1 + a_2) \alpha^j$$

$$\doteq \sum_{j=0}^{\infty} |t| a_1 \alpha^j + \sum_{j=0}^{\infty} |t| (a_1 + a_2) \alpha^j$$

$$\doteq \frac{|t| a_1}{1 - \alpha} + \frac{|t| (a_1 + a_2)}{1 - \alpha} \quad .$$

Since the right hand side of (3.1.3) is finite, $\log_e M(t)$ converges and hence $M(t)$ also converges for $-\infty < t < 0$. Consider the case

$$0 \leq t < \infty \quad , \qquad \text{i.e.} \quad t = |t| \quad .$$

Then $\left|\log_e M(t)\right|$ satisfies

$$(3.1.4) \qquad \left|\log_e M(t)\right| = \sum_j \log_e(\pi_1 e^{|t| a_1\alpha^j} + \pi_2 e^{|t| a_1\alpha^j})$$

$$\leq \sum_j \log_e e^{|t| a_1\alpha^j}$$

$$= \sum_j |t| \ a_1 \ \alpha^j$$

$$= \frac{|t| \ a_1}{1 - \alpha} \quad .$$

Since the right hand side of (3.1.4) is finite, $\log_e M(t)$ converges and hence $M(t)$ also converges for $0 \leq t < \infty$. Thus considering (3.1.3) and (3.1.4) together, $M(t)$ converges for $-\infty < t < \infty$ i.e. for finite values of t .

In the case $\pi_1 = \pi_2 = \frac{1}{2}$, the asymptotic moment generating function for the equal α case is, therefore,

$$(3.1.5) \qquad M(t) = \prod_{j=0}^{\infty} \tfrac{1}{2}(e^{t\alpha^j a_1} + e^{t\alpha^j a_2}) \quad ,$$

and the characteristic function of the asymptotic distribution is by (2.3.6)

$$(3.1.6) \quad \phi(t) = \prod_{j=o}^{\infty} \tfrac{1}{2}(e^{it\alpha^j a_1} + e^{it\alpha^j a_2})$$

where $\quad i^2 = -1 \quad .$

Since by (1.1.26) and (1.1.28) ,

$$a_1 = (1 - \alpha)\lambda_1 \quad \text{and} \quad a_2 = (1 - \alpha)\lambda_2 \quad ,$$

$$(3.1.7) \quad \phi(t) = \prod_{j=o}^{\infty} \tfrac{1}{2}(e^{it(1-\alpha)\lambda_1\alpha^j} + e^{it(1-\alpha)\lambda_2\alpha^j})$$

$$= \prod_{j=o}^{\infty} \tfrac{1}{2}\Big\{\cos\big[(1 - \alpha)\lambda_1 \; \alpha^j \; t\big] + i \sin\big[(1 - \alpha)\lambda_1 \; \alpha^j \; t\big]$$

$$+ \cos\big[(1 - \alpha)\lambda_2 \; \alpha^j \; t\big] + i \sin\big[(1 - \alpha)\lambda_2 \; \alpha^j \; t\big]\Big\}$$

$$= \prod_{j=o}^{\infty} \Big[\cos\frac{(1 - \alpha)(\alpha^j\lambda_1 + \alpha^j\lambda_2)t}{2} \cdot \cos\frac{(1-\alpha)(\alpha^j\lambda_1 - \alpha^j\lambda_2)t}{2}$$

$$+ i \sin\frac{(1 - \alpha)(\alpha^j\lambda_1 + \alpha^j\lambda_2)t}{2} \cdot \cos\frac{(1 - \alpha)(\alpha^j\lambda_1 - \alpha^j\lambda_2)t}{2} \Big]$$

$$= \prod_{j=o}^{\infty} \cos \frac{(1 - \alpha)(\alpha^j\lambda_1 - \alpha^j\lambda_2)t}{2} \cdot \left[\cos \frac{(1 - \alpha)(\alpha^j\lambda_1 + \alpha^j\lambda_2)t}{2} \right.$$

$$\left. + i \sin \frac{(1 - \alpha)(\alpha^j\lambda_1 + \alpha^j\lambda_2)t}{2} \right]$$

$$= \prod_{j=o}^{\infty} e^{\frac{it}{2}(1-\alpha)(\lambda_1+\lambda_2)\alpha^j} \cos \frac{(1 - \alpha)(\alpha^j\lambda_1 - \alpha^j\lambda_2)t}{2}$$

$$= e^{\frac{it(\lambda_1+\lambda_2)}{2}} \prod_{j=o}^{\infty} \cos \frac{(1 - \alpha)(\lambda_1 - \lambda_2)\alpha^j\, t}{2}$$

$$= e^{\frac{it(\lambda_1+\lambda_2)}{2}} \prod_{j=o}^{\infty} \cos \alpha^j \psi$$

where $\psi = (1 - \alpha)(\lambda_1 - \lambda_2)\frac{t}{2}$.

Thus (3.1.7) gives the characteristic function of the asymptotic distribution of the two experimenter-controlled events model with $\alpha_1 = \alpha_2 = \alpha$ and with event E_1 and event E_2 occurring an equal proportion of times. The conditions $\pi_1 = \pi_2 = \frac{1}{2}$ and $\alpha_1 = \alpha_2 = \alpha$, imply that event E_1 and event E_2 occur equally often and that they (e.g. in the case of T-Maze experiment, reward and non-reward) have equal but opposite effects on behaviour. What follows in this section will be solutions of the above characteristic function for some special values of α .

(a) Case I. $\alpha = 0$, $\lambda_1 \neq \lambda_2$, $\pi_1 = \pi_2 = \frac{1}{2}$

i.e. $Q_1\, p = \lambda_1$, $Q_2\, p = \lambda_2$, and the operators are applied an equal proportion of times.

With $\alpha = 0$, the characteristic function of the asymptotic distribution (3.1.7) becomes

$$(3.1.8) \qquad \phi(t) = e^{\frac{it(\lambda_1+\lambda_2)}{2}} \prod_{j=0}^{\infty} \cos \alpha^j (1-\alpha)(\lambda_1 - \lambda_2)\frac{t}{2}$$

$$= e^{\frac{it(\lambda_1+\lambda_2)}{2}} \cos(\lambda_1 - \lambda_2)\frac{t}{2}$$

$$= e^{\frac{it(\lambda_1+\lambda_2)}{2}} \frac{1}{2}\left(e^{\frac{it}{2}(\lambda_1-\lambda_2)} + e^{\frac{it}{2}(\lambda_2-\lambda_1)}\right)$$

$$= \frac{1}{2}\, e^{it\lambda_1} + \frac{1}{2}\, e^{it\lambda_2} \; .$$

Thus the distribution is binomial with variate values λ_1 and λ_2 each with probability $\frac{1}{2}$. Hence if $\lambda_1 = 1$ and $\lambda_2 = 0$, then

$$(3.1.9) \qquad \text{Prob. } (p = 0) = \frac{1}{2}$$

$$\text{and} \qquad \text{Prob. } (p = 1) = \frac{1}{2} \; .$$

If we transform p to $x = 2p - 1$, it is the simple binomial distribution with

$$(3.1.10) \qquad \text{Prob. } (x = -1) = \frac{1}{2}$$

Prob. $(x = 1) = \frac{1}{2}$.

The characteristic function of the asymptotic distribution of x is thus $\phi_x(t) = \cos t$.

(b) Case II. $\alpha = 1$, $\lambda_1 \neq \lambda_2$, $\pi_1 = \pi_2 = \frac{1}{2}$,

i.e. $Q_1 p = p$ $Q_2 p = p$ and the operators are applied an equal proportion of times.

With $\alpha = 1$, the characteristic function of the asymptotic distribution (3.1.7) becomes

$$\text{(3.1.11)} \quad \phi(t) = e^{\frac{it(\lambda_1+\lambda_2)}{2}} \prod_{j=o}^{\infty} \cos \alpha^j (1 - \alpha)(\lambda_1 - \lambda_2)\frac{t}{2}$$

$$= e^{\frac{it(\lambda_1+\lambda_2)}{2}} \quad .$$

That is, the asymptotic distribution is concentrated at the point $p = \frac{\lambda_1 + \lambda_2}{2}$ with

$$\text{(3.1.12)} \quad \text{Prob. } (p = \frac{\lambda_1 + \lambda_2}{2}) = 1 \quad .$$

If $\lambda_1 = 1$ and $\lambda_2 = 0$, then

$$\text{(3.1.13)} \quad \text{Prob. } (p = \tfrac{1}{2}) = 1 \quad .$$

If we transform p to $x = 2p - 1$, the distribution is concentrated at the point $x = 0$ with

(3.1.14) Prob. $(x = 0) = 1$.

(c) Case III. $\alpha = \frac{1}{2}$, $\lambda_1 \neq \lambda_2$, $\pi_1 = \pi_2 = \frac{1}{2}$

i.e. $Q_1\, p = \frac{1}{2} p + (1 - \frac{1}{2})\lambda_1$, $Q_2\, p = \frac{1}{2} p + (1 - \frac{1}{2})\lambda_2$ and the operators are applied an equal proportion of times.

With these values, the characteristic function of the asymptotic distribution (3.1.7) for two experimenter-controlled events becomes

(3.1.15) $$\phi(t) = e^{i(\lambda_1+\lambda_2)\frac{t}{2}} \prod_{j=o}^{\infty} \cos \frac{\psi}{2^j}$$

$$= e^{\frac{1}{2}i(\lambda_1+\lambda_2)t} \cos\psi \prod_{j=1}^{\infty} \cos \frac{\psi}{2^j}$$

$$= e^{\frac{1}{2}i(\lambda_1+\lambda_2)t} \cos\psi \, \frac{\sin\psi}{\psi} \quad \text{(by [5] page 1, 2)}$$

$$= \frac{e^{\frac{1}{2}i(\lambda_1+\lambda_2)t}}{\frac{(\lambda_1-\lambda_2)t}{4}} \cos \frac{(\lambda_1 - \lambda_2)t}{4} \cdot \sin\frac{(\lambda_1 - \lambda_2)t}{4}$$

$$= \frac{e^{\frac{1}{2}i(\lambda_1+\lambda_2)t}}{(\lambda_1 - \lambda_2)\frac{t}{2}} \sin\frac{(\lambda_1 - \lambda_2)t}{2} \quad .$$

By the inversion theorem [4] page 91, the asymptotic probability density function of the distribution of p, the probability of making response A_1 is

$$(3.1.16)\quad f(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{\frac{1}{2}i(\lambda_1+\lambda_2)t}}{(\lambda_1-\lambda_2)\frac{t}{2}} \sin\frac{(\lambda_1-\lambda_2)t}{2} \cdot e^{-itp}\, dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\left[\cos(\lambda_1+\lambda_2-2p)\frac{t}{2} + i\sin(\lambda_1+\lambda_2-2p)\frac{t}{2}\right]\sin(\lambda_1-\lambda_2)\frac{t}{2}}{(\lambda_1-\lambda_2)\frac{t}{2}}\, dt$$

Since $\dfrac{\sin\frac{1}{2}(\lambda_1+\lambda_2-2p)t\,\sin(\lambda_1-\lambda_2)\frac{t}{2}}{(\lambda_1-\lambda_2)\frac{t}{2}}$ is an odd function of t,

$$\frac{i}{2\pi}\int_{-\infty}^{\infty} \frac{\sin\frac{1}{2}(\lambda_1+\lambda_2-2p)t\,\sin(\lambda_1-\lambda_2)\frac{t}{2}}{(\lambda_1-\lambda_2)\frac{t}{2}}\, dt = 0$$

and so

$$(3.1.17)\quad f(p) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\cos\frac{1}{2}(\lambda_1+\lambda_2-2p)t\ \sin\frac{1}{2}(\lambda_1-\lambda_2)t}{\frac{1}{2}(\lambda_1-\lambda_2)t}\, dt$$

$$= \frac{1}{\pi}\int_{0}^{\infty} \frac{\cos\frac{1}{2}(\lambda_1+\lambda_2-2p)t\ \sin\frac{1}{2}(\lambda_1-\lambda_2)t}{\frac{1}{2}(\lambda_1-\lambda_2)t}\, dt \quad .$$

Since the integrand is an even function, let

$$(3.1.18)\qquad y = \tfrac{1}{2}(\lambda_1-\lambda_2)t \quad , \qquad \text{i.e.}$$

$$t = \frac{2y}{\lambda_1 - \lambda_2} \quad \text{and} \quad dt = \frac{2dy}{\lambda_1 - \lambda_2} \quad .$$

Hence

$$(3.1.19) \quad f(p) = \frac{2}{\pi(\lambda_1 - \lambda_2)} \int_0^\infty \cos\left[\frac{(\lambda_1 + \lambda_2 - 2p)y}{\lambda_1 - \lambda_2}\right] \frac{\sin y}{y}\, dy \quad .$$

By [6] page 63 it follows that

$$\begin{aligned} f(p) &= 0 & &\text{if} \quad \frac{\lambda_1 + \lambda_2 - 2p}{\lambda_1 - \lambda_2} < -1 \quad \text{or} > 1 \\ &= \frac{2}{\pi(\lambda_1 - \lambda_2)} \cdot \frac{\pi}{4} & &\text{if} \quad \frac{\lambda_1 + \lambda_2 - 2p}{\lambda_1 - \lambda_2} = \pm 1 \\ &= \frac{2}{\pi(\lambda_1 - \lambda_2)} \cdot \frac{\pi}{2} & &\text{if} \quad -1 < \frac{\lambda_1 + \lambda_2 - 2p}{\lambda_1 - \lambda_2} < 1 \quad . \end{aligned}$$

Thus

$$\begin{aligned} (3.1.20) \quad f(p) &= 0 & &\text{if} \quad p > \lambda_1 \quad \text{or} \quad p > \lambda_2 \\ &= \frac{1}{2(\lambda_1 - \lambda_2)} & &\text{if} \quad p = \lambda_1 \quad \text{or} \quad p = \lambda_2 \\ &= \frac{1}{\lambda_1 - \lambda_2} & &\text{if} \quad \lambda_2 < p < \lambda_1 \quad . \end{aligned}$$

This is shown by figure 3.1 .

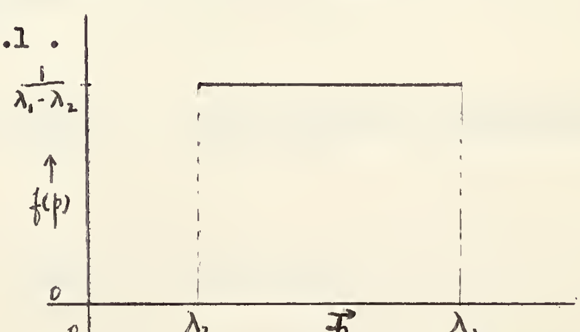

Fig. 3.1 Asymptotic distribution of p for two experimenter-controlled events with $\alpha_1 = \alpha_2 = \alpha = \frac{1}{2}$ and $\lambda_1 > \lambda_2$.

Hence the asymptotic distribution of p (for $\alpha_1 = \alpha_2 = \frac{1}{2}$, $\lambda_1 \neq \lambda_2$ and $\pi_1 = \pi_2 = \frac{1}{2}$) is the uniform distribution $f(p) = \dfrac{1}{\lambda_1 - \lambda_2}$ lying between the limit points λ_2 and λ_1. If $\lambda_2 = 0$, and $\lambda_1 = 1$, the asymptotic distribution of p lies uniformly between 0 and 1. That is

$$
\begin{aligned}
(3.1.21)\quad f(p) &= 1 && \text{for} \quad 0 < p < 1 \\
&= 0 && \text{for} \quad p < 0 \text{ and } p > 1 .
\end{aligned}
$$

This is shown in figure 3.2. This result was obtained independently by Karlin [2] using entirely different arguments.

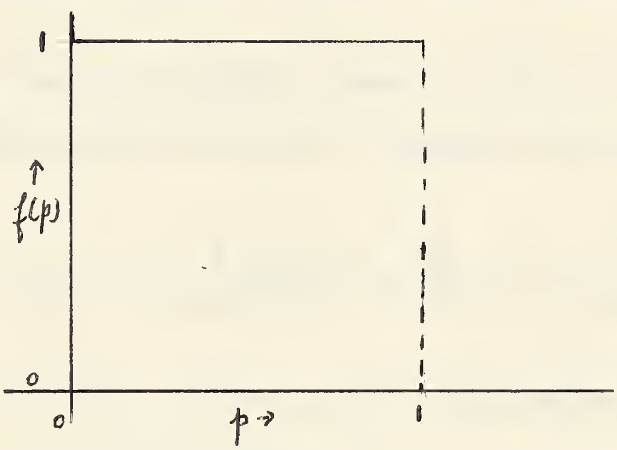

Fig. 3.2 Asymptotic distribution of p for two experimenter-controlled events with $\alpha_1 = \alpha_2 = \frac{1}{2}$ and $\lambda_1 = 1$, $\lambda_2 = 0$.

If we transform p to $x = 2p - 1$, the asymptotic distribution of x is

$$
\begin{aligned}
(3.1.22)\quad f(x) &= \tfrac{1}{2} && \text{for} \quad -1 < x < 1 \\
&= 0 && \text{for} \quad x < -1 \text{ and } x > 1 .
\end{aligned}
$$

This is shown in figure 3.2 .

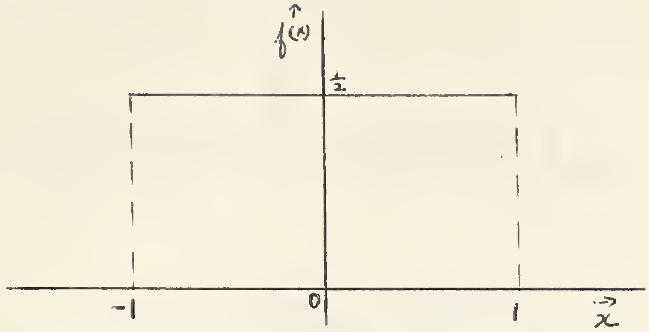

Fig. 3.3 Asymptotic distribution of $x = 2p - 1$ for two experimenter-controlled events with $\alpha_1 = \alpha_2 = \frac{1}{2}$ and $\lambda_1 = 1$, $\lambda_2 = 0$.

As a check on these results we compare some asymptotic moments calculated from the recurrence formula (1.2.7) and from the density function (3.1.20) .

From the recurrence formula (1.2.7) for two experimenter-controlled events of chapter one, in general

$$V_{m,n+1} = \pi_1 \sum_{u=o}^{m} \binom{m}{u} a_1^{m-u} \alpha_1^{u} V_{u,n} + \pi_2 \sum_{u=o}^{m} \binom{m}{u} a_2^{m-u} \alpha_2^{u} V_{u,n} .$$

When $\pi_1 = \pi_2 = \frac{1}{2}$, $\alpha_1 = \alpha_2 = \frac{1}{2}$, the recurrence formula becomes

$$(3.1.23) \quad V_{m,n+1} = \sum_{u=o}^{m} \binom{m}{u} (a_1^{m-u} + a_2^{m-u}) (\tfrac{1}{2})^{u+1} V_{u,n} .$$

Letting n approach infinity,

$$(3.1.24) \quad \lim_{n \to \infty} V_{m,n+1} = V_{m,\infty}$$

$$= \sum_{u=0}^{m} \binom{m}{u} (a_1^{m-u} + a_2^{m-u}) (\tfrac{1}{2})^{u+1} V_{u,\infty} \quad .$$

Thus

$$(3.1.25) \quad V_{1,\infty} = \sum_{u=0}^{1} \binom{1}{u} (a_1^{1-u} + a_2^{1-u}) (\tfrac{1}{2})^{u+1} V_{u,\infty}$$

$$= (a_1 + a_2)\tfrac{1}{2} V_{0,\infty} + 2(\tfrac{1}{2})^2 V_{1,\infty}$$

$$= \tfrac{1}{2}(a_1 + a_2) + \tfrac{1}{2} V_{1,\infty} \quad ,$$

i.e. $\quad V_{1,\infty} = a_1 + a_2$

$$= \frac{\lambda_1 + \lambda_2}{2} \quad .$$

From the asymptotic density function (3.1.20)

$$(3.1.26) \quad V_{1,\infty} = \int_{\lambda_2}^{\lambda_1} \frac{p}{\lambda_1 - \lambda_2}\, dp$$

$$= \frac{(\lambda_1^2 - \lambda_2^2)}{2(\lambda_1 - \lambda_2)}$$

$$= \frac{\lambda_1 + \lambda_2}{2}$$

$$= a_1 + a_2 \quad .$$

Thus the asymptotic first moment calculated from the recurrence formula is the same as that calculated from the density function.

As to the second moment, we have from the recurrence formula,

$$(3.1.27)\quad V_{2,\infty} = \sum_{u=0}^{2} \binom{2}{u}(a_1^{2-u} + a_2^{2-u})(\tfrac{1}{2})^{u+1}\, V_{u,\infty}$$

$$= (a_1^2 + a_2^2)(\tfrac{1}{2})V_{0,\infty} + 2(a_1 + a_2)(\tfrac{1}{2})^2 V_{1,\infty} + 2(\tfrac{1}{2})^3\, V_{2,\infty}$$

$$= \frac{2}{3}\left[\ (a_1^2 + a_2^2) + (a_1 + a_2)^2\ \right]$$

$$= \frac{4}{3}(a_1^2 + a_1\, a_2 + a_2^2)$$

$$= \frac{1}{3}(\lambda_1^2 + \lambda_1 \lambda_2 + \lambda_2^2) \quad .$$

From the density function

$$(3.1.28)\quad V_{2,\infty} = \int_{\lambda_2}^{\lambda_1} \frac{p^2}{\lambda_1 - \lambda_2}\, dp$$

$$= \frac{\lambda_1^3 - \lambda_2^3}{3(\lambda_1 - \lambda_2)}$$

$$= \frac{1}{3}(\lambda_1^2 + \lambda_1 \lambda_2 + \lambda_2^2) \quad .$$

Thus the results also agree in the case of the second moment.

In general, from the density function,

$$(3.1.29) \quad V_{m,\infty} = \int_{\lambda_2}^{\lambda_1} \frac{p^m}{\lambda_1 - \lambda_2}\, dp$$

$$= \frac{\lambda_1^{m+1} - \lambda_2^{m+1}}{(m+1)(\lambda_1 - \lambda_2)}$$

$$= \frac{2^m(a_1^{m+1} - a_2^{m+1})}{(m+1)(a_1 - a_2)} \quad ,$$

and from the recurrence formula,

$$V_{m,\infty} = \frac{1}{2} \sum_{u=0}^{m} \binom{m}{u} (a_1^{m-u} + a_2^{m-u}) \left(\frac{1}{2}\right)^u V_{u,\infty} \quad .$$

It has been verified that for $m = 1, 2,$ the results obtained from the recurrence formula are equal to those obtained from the density function. Assume that this holds for the mth case. Then for the $(m + 1)$th case from the recurrence formula

$$(3.1.30) \quad V_{m+1,\infty} = \frac{1}{2} \sum_{u=0}^{m+1} \binom{m+1}{u} (a_1^{m+1-u} + a_2^{m+1-u}) \left(\frac{1}{2}\right)^u V_{u,\infty}$$

$$= \frac{1}{2} \sum_{u=0}^{m} \binom{m+1}{u} (a_1^{m+1-u} + a_2^{m+1-u}) \left(\frac{1}{2}\right)^u V_{u,\infty}$$

$$+ \frac{1}{2}(1 + 1)\left(\frac{1}{2}\right)^{m+1} V_{m+1,\infty} \quad ,$$

i.e.

$$\left[1 - (\tfrac{1}{2})^{m+1}\right] V_{m+1,\infty} = \frac{1}{2} \sum_{u=o}^{m} \binom{m+1}{u} (a_1^{m+1-u} + a_2^{m+1-u}) \frac{(\frac{1}{2})^u \, 2^u (a_1^{u+1} - a_2^{u+1})}{(u+1)(a_1 - a_2)}$$

$$= \frac{1}{2} \sum_{u=o}^{m} \binom{m+1}{u} \frac{(a_1^{m+2} - a_2^{m+2})}{(u+1)(a_1 - a_2)}$$

$$+ \frac{1}{2} \sum_{u=o}^{m} \binom{m+1}{u} \frac{(a_1^{u+1} \, a_2^{m+1-u} - a_1^{m+1-u} \, a_2^{u+1})}{(u+1)(a_1 - a_2)}$$

$$= \frac{1}{2} \frac{(a_1^{m+2} - a_2^{m+2})}{(a_1 - a_2)} \left[\sum_{u=o}^{m+1} \binom{m+1}{u} \frac{1}{u+1} - \frac{1}{m+2} \right]$$

$$+ \sum_{u=o}^{m} \binom{m+1}{u} \frac{a_1^{u+1} \, a_2^{m+1-u}}{(u+1)(a_1 - a_2)} - \sum_{u=o}^{m} \binom{m+1}{u} \frac{a^{m+1-u} \, a^{u+1}}{(u+1)(a_1 - a_2)}$$

$$= \frac{1}{2} \frac{(a_1^{m+2} - a_2^{m+2})}{(a_1 - a_2)} \left[\sum_{u=o}^{m+1} \binom{m+1}{u} \frac{1}{u+1} - \frac{1}{m+2} \right]$$

$$+ \sum_{u=o}^{m} \binom{m+1}{u} \frac{a_1^{u+1} \, a_2^{m+1-u}}{(u+1)(a_1 - a_2)} - \sum_{u=o}^{m} \binom{m+1}{m-u} \frac{a_1^{m-u+1} \, a_2^{m+1-(m-u)}}{\left[(m-u) + 1\right](a_1 - a_2)} \; .$$

In the last term, let $v = m - u$. Then when $u = o$, $v = m$ and when $u = m$, $v = o$. Thus we get the term

$\sum_{v=o}^{m} \binom{m+1}{v} \frac{a_1^{v+1} \, a_2^{m+1-v}}{v+1}$ which is the same as the second term and

so they cancel to 0 . Thus

$$\left[1 - (\tfrac{1}{2})^{m+1}\right] V_{m+1,\infty} = \frac{1}{2}\,\frac{a_1^{m+2} - a_2^{m+2}}{a_1 - a_2}\left[\sum_{u=0}^{m+1}\binom{m+1}{u}\frac{1}{u+1} - \frac{1}{m+2}\right] .$$

To see that

$$\sum_{u=0}^{m+1}\binom{m+1}{u}\frac{1}{u+1} = \frac{2^{m+2} - 1}{m+2} ,$$

consider $\int_0^1 (1 + x)^{m+1}\,dx$. Expanding the integrand by the binomial theorem, we obtain

(3.1.31) $$\int_0^1 (1 + x)^{m+1}\,dx = \int_0^1 \sum_{u=0}^{m+1}\binom{m+1}{u} x^u\,dx$$

$$\left[\frac{(1 + x)^{m+2}}{m+2}\right]_{x=0}^{x=1} = \left[\sum_{u=0}^{m+1}\binom{m+1}{u}\frac{x^{u+1}}{u+1}\right]_{x=0}^{x=1}$$

or

$$\frac{2^{m+2} - 1}{m+2} = \sum_{u=0}^{m}\binom{m+1}{u}\frac{1}{u+1} .$$

Hence

$$(1 - \frac{1}{2^{m+1}})V_{m+1,\infty} = \frac{1}{2}\,\frac{a_1^{m+2} - a_2^{m+2}}{a_1 - a_2}\left[\frac{2^{m+2} - 1}{m+2} - \frac{1}{m+2}\right]$$

and

(3.1.32) $$V_{m+1,\infty} = \frac{1}{2}\,\frac{a_1^{m+2} - a_2^{m+2}}{a_1 - a_2} \cdot \frac{2^{m+2} - 2}{m+2} \cdot \frac{2^{m+1}}{2^{m+1} - 1}$$

$$= \frac{a_1^{m+2} - a_2^{m+2}}{a_1 - a_2} \cdot \frac{2^{m+1} - 1}{m + 2} \cdot \frac{2^{m+1}}{2^{m+1} - 1}$$

$$= \frac{2^{m+1}}{m + 2} \cdot \frac{a_1^{m+2} - a_2^{m+2}}{a_1 - a_2}$$

$$= \frac{\lambda_1^{m+2} - \lambda_2^{m+2}}{(m + 2)(\lambda_1 - \lambda_2)}$$

$$= \int_{\lambda_2}^{\lambda_1} \frac{p^{m+1}}{\lambda_1 - \lambda_2} \, dp \quad .$$

Thus in general, the moments obtained from the recurrence formula (1.2.7) are the same as those calculated from the density function (3.1.20) which we derived.

(d) Case IV. $\alpha = \frac{1}{4}$, $\lambda_1 = 1$, $\lambda_2 = 0$ and $\pi_1 = \pi_2 = \frac{1}{2}$,

i.e. $Q_1\, p = \frac{1}{4}p + (1 - \frac{1}{4})$, $Q_2\, p = \frac{1}{4}p$ and both operators are applied equally often.

For this case, the characteristic function of the asymptotic distribution (3.1.7) becomes

$$(3.1.33) \quad \phi(t) = e^{it(\lambda_1 + \lambda_2)/2} \prod_{j=0}^{\infty} \cos \alpha^j \psi$$

$$= e^{it/2} \prod_{j=0}^{\infty} \cos \frac{\psi}{2^{2j}}$$

where $\qquad \psi = (1 - \alpha)(\lambda_1 - \lambda_2)\frac{t}{2} = \frac{3t}{8}$.

The interval of the variable p , the asymptotic probability of making response A_1 is given by the limit points λ_1 and λ_2 . In this case the interval is $[1,0]$. For convenience, it is changed to $[1,-1]$ by introducing the variable $x = 2p - 1$. The characteristic function of the asymptotic distribution of x is

$$(3.1.34) \quad \phi_x(t) = E(e^{itx}) = \int_0^1 f(p)\, e^{itx}\, dp$$

$$= \int_0^1 f(p)\, e^{it(2p-1)}\, dp$$

$$= e^{-it} \int_0^1 f(p)\, e^{2itp}\, dp$$

$$= e^{-it}\, \phi(2t)$$

$$= e^{-it} \cdot e^{it} \prod_{j=0}^{\infty} \cos \frac{(1-\alpha)(\lambda_1 - \lambda_2)t}{2^{2j}}$$

$$= \prod_{j=0}^{\infty} \cos \frac{3t}{2^{2j+2}}$$

$$= \frac{\prod_{j=0}^{\infty} \cos \frac{3t}{2^{2j+2}} \cdot \prod_{j=0}^{\infty} \cos \frac{3t}{2^{2j+1}}}{\prod_{j=0}^{\infty} \cos \frac{3t}{2^{2j+1}}}$$

$$= \frac{\prod_{j=1}^{\infty} \cos \frac{3t}{2^j}}{\prod_{j=o}^{\infty} \cos \frac{3t}{2^{2j+1}}} \quad .$$

Since by [5] page 1, 2

$$(3.1.35) \quad \prod_{j=1}^{\infty} \cos \frac{3t}{2^j} = \frac{\sin 3t}{3t} \quad ,$$

$$(3.1.36) \quad \phi_x(t) = \frac{\sin 3t}{3t} \prod_{j=o}^{\infty} \sec \frac{3t}{2^{2j+1}} \quad .$$

Thus the characteristic function of the asymptotic distribution of x is the term $\frac{\sin 3t}{3t}$ modified by an infinite product $\prod_{j=o}^{\infty} \sec \frac{3t}{2^{2j+1}}$. Let

$$(3.1.37) \quad \phi_o(t) = \frac{\sin 3t}{3t} \quad ,$$

$$\phi_1(t) = \frac{\sin 3t}{3t} \cdot \sec \frac{3t}{2}$$

$$= \frac{\sin \frac{3t}{2}}{\frac{3t}{2}} \quad ,$$

$$\phi_2(t) = \frac{\sin 3t}{3t} \sec \frac{3t}{2} \cdot \sec \frac{3t}{2^3}$$

$$= \frac{\sin \frac{3t}{2}}{\frac{3t}{2}} \cdot \sec \frac{3t}{2^3}$$

$$= \frac{\sin \frac{3t}{8}}{\frac{3t}{8}} \cdot \cos \frac{3t}{2^{2}}$$

$$= \phi_1\left(\frac{t}{4}\right) \cos \frac{3t}{4} \quad ,$$

$$\phi_3(t) = \frac{\sin 3t}{3t} \cdot \sec \frac{3t}{2} \cdot \sec \frac{3t}{2^3} \cdot \sec \frac{3t}{2^5}$$

$$= \frac{\sin \frac{3t}{2^3}}{\frac{3t}{2^3}} \cdot \cos \frac{3t}{4} \cdot \sec \frac{3t}{2^5}$$

$$= \sin \frac{3t}{2^5} \cdot \cos \frac{3t}{4} \cdot \cos \frac{3t}{2^4}$$

$$= \phi_2\left(\frac{t}{4}\right) \cdot \cos \frac{3t}{4} \quad ,$$

and in general let

$$\phi_n(t) = \frac{\sin 3t}{3t} \cdot \prod_{j=1}^{n} \sec \left(\frac{3t}{2^{2j-1}} \right) \quad \text{where } n \geq 1 \; . \tag{3.1.38}$$

Furthermore assume that for the nth case ($n \geq 2$) ,

$$\phi_n(t) = \phi_{n-1}\left(\frac{t}{4}\right) \cdot \cos \frac{3t}{4} \tag{3.1.39}$$

$$= \frac{\sin \frac{3t}{2^{2n-1}}}{\frac{3t}{2^{2n-1}}} \cdot \prod_{j=1}^{n-1} \cos \left(\frac{3t}{2^{2j}} \right) \; .$$

Then for the $(n + 1)$th case, where $n \geq 2$

$$(3.1.40)\quad \phi_{n+1}(t) = \frac{\sin 3t}{3t} \prod_{j=1}^{n+1} \sec\left(\frac{3t}{2^{2j-1}}\right)$$

$$= \phi_n(t) \sec\left(\frac{3t}{2^{2n+1}}\right)$$

$$= \frac{\sin \frac{3t}{2^{2n-1}}}{\frac{3t}{2^{2n-1}}} \prod_{j=1}^{n-1} \cos \frac{3t}{2^{2j}} \cdot \sec\left(\frac{3t}{2^{2n+1}}\right)$$

$$= \frac{\sin \frac{3t}{2^{2n+1}}}{\frac{3t}{2^{2n+1}}} \prod_{j=1}^{n} \cos \frac{3t}{2^{2j}}$$

$$= \phi_n\left(\frac{t}{4}\right) \cos \frac{3t}{4} \;,$$

which is of the same form as in the nth case. Thus in general for $n \geq 2$,

$$(3.1.41)\qquad \phi_n(t) = \phi_{n-1}\left(\frac{t}{4}\right) \cos \frac{3t}{4} \;, \quad \text{and}$$

$$\phi_x(t) = \lim_{n \to \infty} \phi_n(t) = \frac{\sin 3t}{3t} \prod_{j=1}^{\infty} \sec\left(\frac{3t}{2^{2j-1}}\right) .$$

By the inversion theorem [4] page 91, the inversion of $\phi_n(t)$ into its density function is

$$(3.1.42)\quad \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_n(t)e^{-itx}\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi_{n-1}\left(\frac{t}{4}\right)\cos\frac{3t}{4}\,e^{-itx}\,dt$$

$$= g_n(x)\ ,\ \text{say}\ .$$

Then by [7] page 63 ,

$$(3.1.43)\quad g_n(x) = 2\left[g_{n-1}(4x-3) + g_{n-1}(4x+3)\right] .$$

By [6] page 63 ,

$$(3.1.44)\quad g_0(x) = \frac{1}{6} \qquad \text{for} \quad -3 < x < 3 ,$$

and $\qquad = 0 \qquad$ for $\quad x < -3$ and $x > 3$,

$$(3.1.45)\quad g_1(x) = \frac{1}{3} \qquad \text{for} \quad -\frac{3}{2} < x < \frac{3}{2} ,$$

$$= 0 \qquad \text{for} \quad x < -\frac{3}{2} \text{ and } x > \frac{3}{2} .$$

By (3.1.43)

$$(3.1.46)\quad g_2(x) = \frac{2}{3} \qquad \text{for} \quad -\frac{9}{8} < x < -\frac{3}{8}\ ,\ \frac{3}{8} < x < \frac{9}{8} ,$$

$$= 0 \qquad \text{for} \quad x < -\frac{9}{8}\ ,\ -\frac{3}{8} < x < \frac{3}{8} \text{ and } x > \frac{9}{8} ,$$

$$(3.1.47)\quad g_3(x) = \frac{4}{3} \qquad \text{for} \quad -\frac{33}{32} < x < -\frac{27}{32}\ ,\ -\frac{21}{32} < x < -\frac{15}{32} ,$$

$$\frac{15}{32} < x < \frac{21}{32}\ ,\ \frac{27}{32} < x < \frac{33}{32} ,$$

$$= 0 \qquad \text{for} \quad x < -\frac{33}{32}\ ,\ -\frac{27}{32} < x < -\frac{21}{32} ,$$

$$-\frac{15}{32} < x < \frac{15}{32}\ ,\ \frac{21}{32} < x < \frac{27}{32} ,$$

$$\text{and} \quad x > \frac{33}{32} ,$$

and so forth.

At the nth stage, there are 2^{n-1} intervals of non-zero probability density. Let them be

(3.1.48) $I_{n,k}$ where $k = 1, \dots 2^{n-1}$, and $n \geq 1$.

For example, for $n = 3$, there are four intervals

(3.1.49) $I_{3,1} : (-\frac{33}{32}, -\frac{27}{32})$ $\quad I_{3,3} : (\frac{15}{32}, \frac{21}{32})$

$I_{3,2} : (-\frac{21}{32}, -\frac{15}{32})$ $\quad I_{3,4} : (\frac{27}{32}, \frac{33}{32})$.

Each such interval $I_{n,k}$ at stage n generates two new intervals at stage $(n + 1)$ in the following manner: divide $I_{n,k}$ into eight equal subintervals $i_{nk,1}, \dots i_{nk,8}$ from left to right; discard $i_{nk,1}$, $i_{nk,4}$, $i_{nk,5}$ and $i_{nk,8}$. Then the two new intervals consist of the union of $i_{nk,2}$, $i_{nk,3}$ and the union of $i_{nk,6}$, and $i_{nk,7}$. The probability distribution at the $(n + 1)$th stage is then obtained by constructing, over the new intervals, identical rectangles of height double that of the preceding stages as shown in figure 3.4 . From this construction it follows that if W_n and h_n are respectively the width and height of a rectangle at the nth stage, then

(3.1.50) $$W_n = \frac{3}{2^{2(n-1)}} \qquad n \geq 1$$

(3.1.51) $$h_n = \frac{2^{n-1}}{3} \qquad n \geq 1 \quad .$$

The total width of the intervals at stage n is

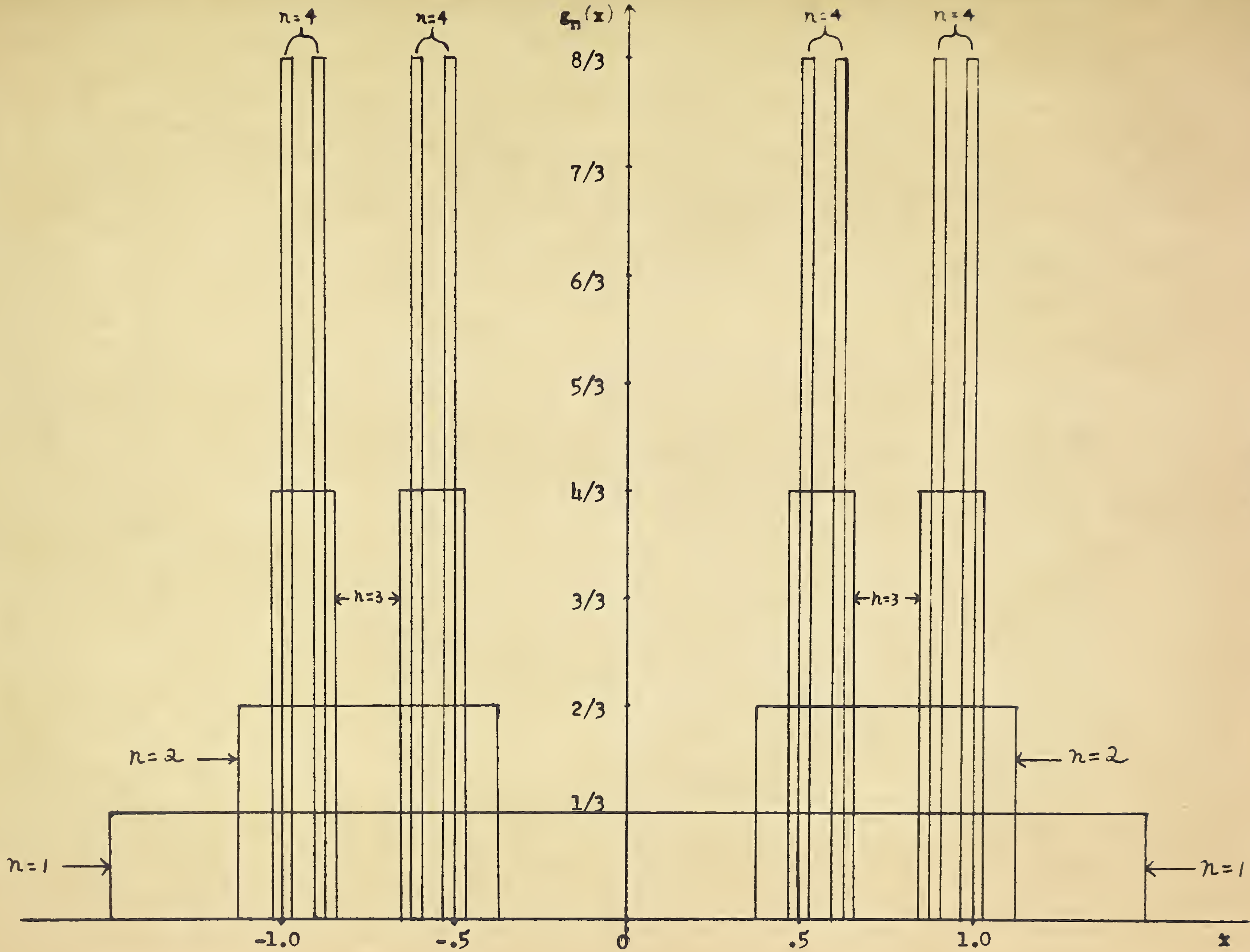

Fig. 3.4 Distribution $g_n(x)$ for $n = 1, 2, 3$, and 4 .

(3.1.52) $W_n = \dfrac{3}{2^{2(n-1)}} \cdot 2^{n-1} = \dfrac{3}{2^{n-1}}$.

Thus as n approaches infinity, W_n approaches 0 , and hence the probability is concentrated on a set of points of Lebesgue measure 0 . The close analogy between the construction described above and that of the Cantor Ternary set suggests that the asymptotic distribution is concentrated on a nonenumerable set of points of Lebesgue measure zero and that the cumulative distribution function is a continuous increasing function which is not an integral ([8] page 329 and page 366). Figure 3.5 shows the cumulative distribution $G_n(x)$ for $n = 3$ and $n = 4$.

(e) Case V: $\alpha = \dfrac{1}{\sqrt{2}}$, $\lambda_2 = 0$, $\lambda_1 = 1$, $\pi_1 = \pi_2 = \frac{1}{2}$.

That is $Q_1\, p = \dfrac{1}{\sqrt{2}}\, p + (1 - \dfrac{1}{\sqrt{2}})$ and $Q_2\, p = \dfrac{1}{\sqrt{2}}\, p$, and both operations occur equally often.

With $\alpha = \dfrac{1}{\sqrt{2}}$, the characteristic function of the asymptotic distribution (3.1.7) becomes

$$(3.1.53)\quad \phi(t) = e^{\frac{it(\lambda_1+\lambda_2)}{2}} \prod_{j=0}^{\infty} \cos \frac{\alpha^j(1-\alpha)(\lambda_1-\lambda_2)t}{2}$$

$$= e^{it/2} \prod_{j=0}^{\infty} \cos \frac{(1-\frac{1}{\sqrt{2}})t}{2^{j/2+1}} \quad .$$

Using the same transformation as in the previous case, let $x = 2p - 1$. The characteristic function of the asymptotic distribution of x following the same procedure as in (3.1.34) is

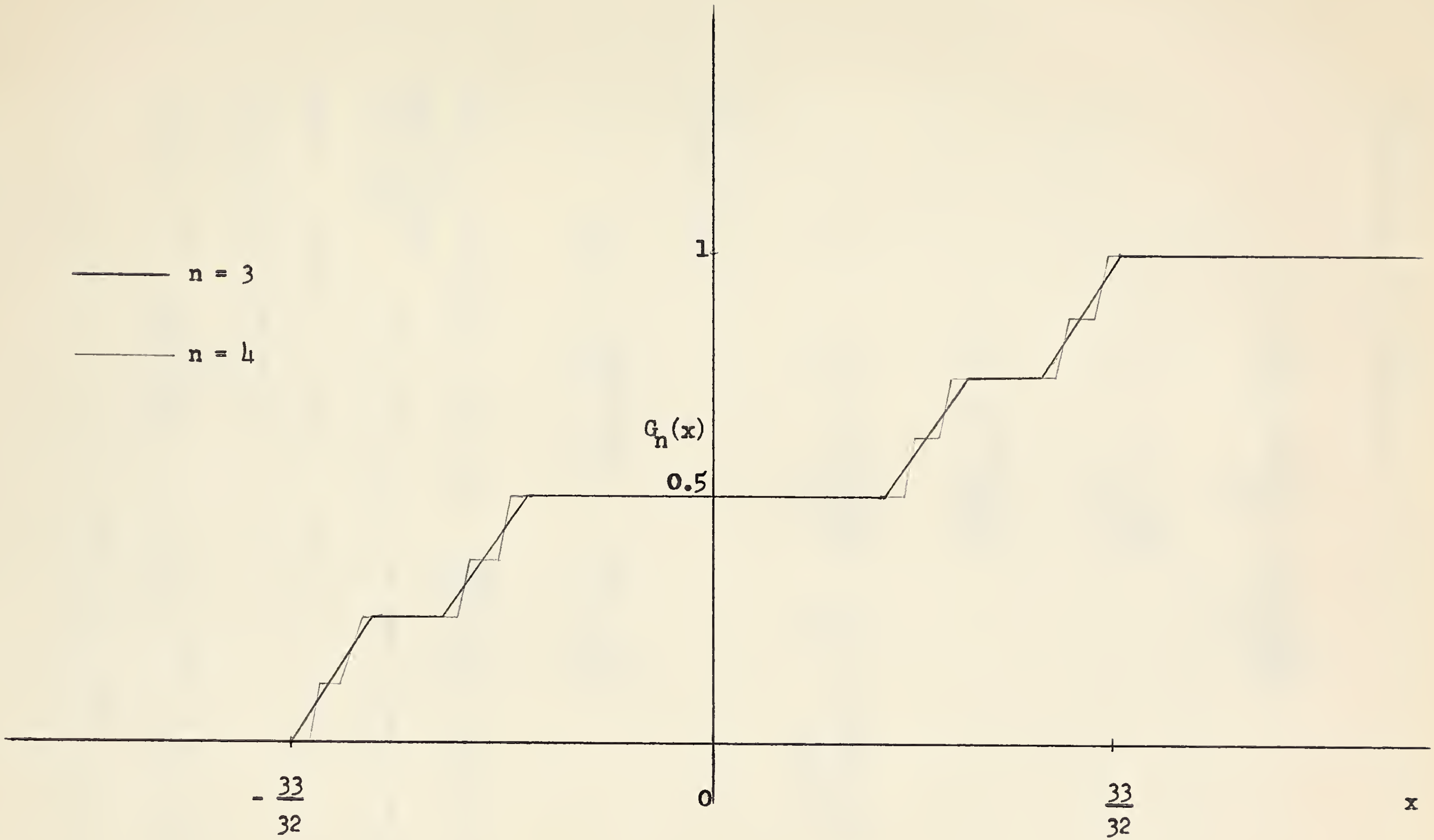

Fig. 3.5. Cumulative distribution $G_n(x)$ for $n = 3$ and 4 .

$$(3.1.54) \quad \phi_x(t) = e^{-it}\,\phi(2t)$$

$$= e^{-it} \cdot e^{it} \prod_{j=o}^{\infty} \cos \frac{(1 - \frac{1}{\sqrt{2}})t}{2^{j/2}}$$

$$= \prod_{j=o}^{\infty} \cos \frac{(1 - \frac{1}{\sqrt{2}})t}{2^{j/2}}$$

$$= \prod_{j=o}^{\infty} \cos \frac{(\sqrt{2} - 1)t}{2^{(j+1)/2}}$$

$$= \prod_{j=o}^{\infty} \cos \frac{(\sqrt{2} - 1)t}{(\sqrt{2})^{2j+2}} \cdot \prod_{j=o}^{\infty} \cos \frac{(\sqrt{2} - 1)t}{(\sqrt{2})^{2j+1}}$$

and by [5] page 1, 2 ,

$$\phi_x(t) = \frac{\sin(\sqrt{2} - 1)t}{(\sqrt{2} - 1)t} \cdot \frac{\sin(2 - \sqrt{2})t}{(2 - \sqrt{2})t}$$

which is the product of the term $\dfrac{\sin(\sqrt{2} - 1)t}{(\sqrt{2} - 1)t}$ and the term $\dfrac{\sin(2 - \sqrt{2})t}{(2 - \sqrt{2})t}$. The first term, by the same inversion method as used in (3.1.16) , (3.1.17) , (3.1.18) , (3.1.19) , is the characteristic function of

$$(3.1.55) \quad f(x) = \frac{1}{2(\sqrt{2} - 1)} \qquad \text{for} \quad -(\sqrt{2} - 1) \leq x \leq (\sqrt{2} - 1)$$

$$= 0 \qquad \text{for} \quad x < -(\sqrt{2} - 1)\ , \quad x > (\sqrt{2} - 1)\ .$$

Similarly, the second term is the characteristic function of

$$(3.1.56) \quad f(x) = \frac{1}{2\sqrt{2}(\sqrt{2}-1)} \qquad \text{for} \quad -\sqrt{2}(\sqrt{2}-1) < x < \sqrt{2}(\sqrt{2}-1)$$

$$= 0 \qquad \text{for} \quad x < -\sqrt{2}(\sqrt{2}-1) \text{ and } x > \sqrt{2}(\sqrt{2}-1) \; .$$

Thus $\phi_x(t)$ is the characteristic function of the convolution ([9] page 250) of the distributions (3.1.55) and (3.1.56). To invert (3.1.54) into the distribution function, consider the more general case

$$(3.1.57) \quad \psi(t) = \frac{\sin at}{at} \cdot \frac{\sin bt}{bt} \qquad \text{where} \quad a > b \; .$$

This is the characteristic function of the convolution of the distributions

$$(3.1.58) \quad g_a(y) = \frac{1}{2a} \qquad \text{for} \quad -a < y < a$$

$$= 0 \qquad \text{for} \quad y < -a \text{ and } y > a \; ,$$

and

$$(3.1.59) \quad g_b(y) = \frac{1}{2b} \qquad \text{for} \quad -b < y < b$$

$$= 0 \qquad \text{for} \quad y < -b \text{ and } y > b \; .$$

The convolution of $g_a(y)$ and $g_b(y)$ is

$$(3.1.60) \quad f(x) = \int_{-\infty}^{\infty} g_a(y)\, g_b(x-y)\,dy \; .$$

To obtain the limits of integration figure 3.6 is drawn.

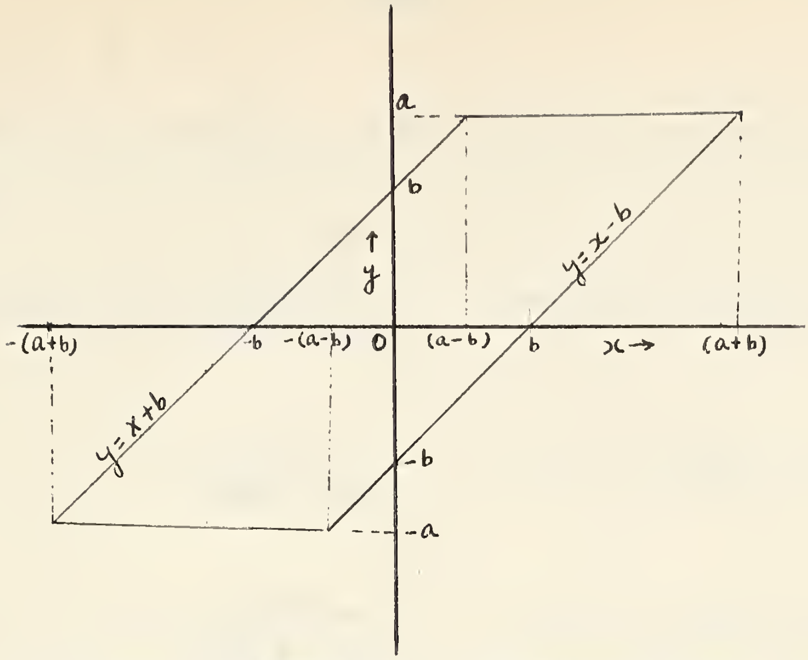

Fig. 3.6 Limits of integration of the convolution of $g_a(y)$ and $g_b(y)$ where $g_a(y) = \frac{1}{2a}$ for $-a \leq y \leq a$, $g_a(y) = 0$ otherwise and $g_b(y) = \frac{1}{2b}$ for $-b \leq y \leq b$, $g_b(y) = 0$, otherwise.

From figure 3.6

$$(3.1.61)\quad f(x) = \frac{1}{2a} \int_{-a}^{x+b} \frac{1}{2b}\, dy = \frac{x + a + b}{4ab} \quad \text{for} \quad -(a + b) < x < -(a - b) ,$$

$$= \frac{1}{4ab} \int_{x-b}^{x+b} dy = \frac{1}{2a} \quad \text{for} \quad -(a - b) < x < (a - b) ,$$

$$= \frac{1}{4ab} \int_{x-b}^{a} dy = \frac{a + b - x}{4ab} \quad \text{for} \quad (a - b) < x < (a + b) ,$$

$$= 0 \quad \text{for} \quad x < -(a + b) \quad \text{and} \quad x > (a + b) .$$

In the characteristic function (3.1.54) considered above,

$a = 2 - \sqrt{2}$, $b = \sqrt{2} - 1$, and $a + b = 1$, $a - b = 3 - 2\sqrt{2}$.

Thus

$$f(x) = \frac{1 + x}{4\sqrt{2}(\sqrt{2} - 1)^2} \qquad \text{for} \quad -1 < x < -(3 - 2\sqrt{2}) \tag{3.1.62}$$

$$= \frac{1}{2\sqrt{2}(\sqrt{2} - 1)} \qquad \text{for} \quad -(3 - 2\sqrt{2}) < x < 3 - 2\sqrt{2}$$

$$= \frac{1 - x}{4\sqrt{2}(\sqrt{2} - 1)^2} \qquad \text{for} \quad 3 - 2\sqrt{2} < x < 1$$

$$= 0 \qquad \text{for} \quad x < -1 \text{ and } x > 1 .$$

This distribution is shown in figure 3.7 .

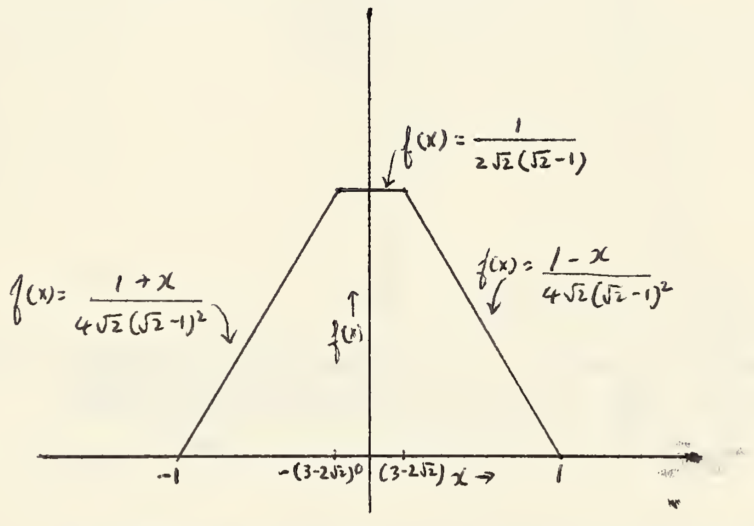

Fig. 3.7. Asymptotic distribution of $x = 2p - 1$ for two experimenter-controlled events with $\alpha_1 = \alpha_2 = \frac{1}{\sqrt{2}}$, $\lambda_2 = 0$, $\lambda_1 = 1$ and $\pi_1 = \pi_2 = \frac{1}{2}$.

Since $x = 2p - 1$, the corresponding asymptotic distribution of p is

$$(3.1.63)\quad f(p) = \frac{p}{\sqrt{2}(\sqrt{2}-1)^2} \qquad \text{for } 0 \le p < \sqrt{2} - 1$$

$$= \frac{1}{\sqrt{2}(\sqrt{2}-1)} \qquad \text{for } \sqrt{2} - 1 \le p < 2 - \sqrt{2}$$

$$= \frac{1-p}{\sqrt{2}(\sqrt{2}-1)^2} \qquad \text{for } 2 - \sqrt{2} \le p < 1 \quad .$$

This is shown in figure 3.8 .

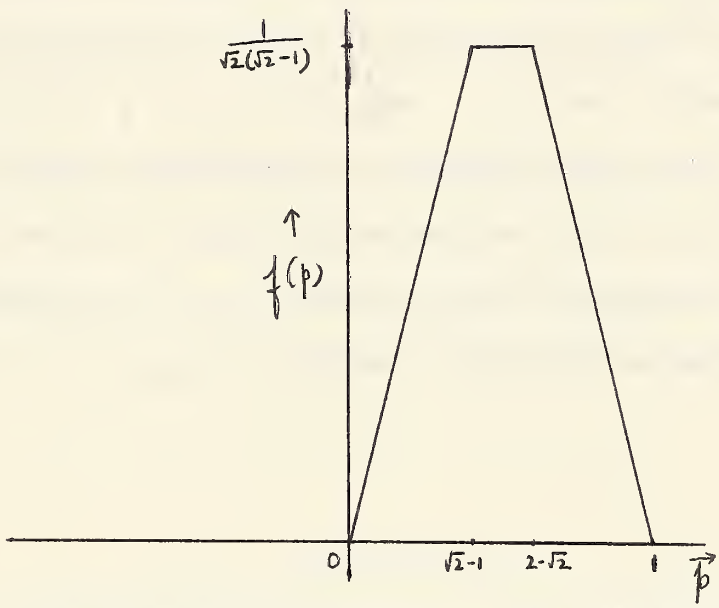

Fig. 3.8. Asymptotic distribution of p from two experimenter-controlled events with $\alpha_1 = \alpha_2 = \frac{1}{\sqrt{2}}$, $\lambda_2 = 0$, $\lambda_1 = 1$ and $\pi_1 = \pi_2 = \frac{1}{2}$.

To summarize this section, let us recapitulate the evolution of the distribution of $x = 2p - 1$ as α ranges from 0 to 1 with $\lambda_1 = 1$ and $\lambda_2 = 0$. For $\alpha = 0$, the distribution is binomial at $x = -1$ and at $x = +1$ each with probability $\frac{1}{2}$.

For $\alpha = \frac{1}{4}$, the distribution is concentrated on a nonenumerable set of points of Lebesgue measure zero. For $\alpha = \frac{1}{2}$, the distribution is uniform from -1 to $+1$ with density function $f(x) = \frac{1}{2}$. For $\alpha = \frac{1}{\sqrt{2}}$, the distribution density function increases in value as x increases from -1 to $-(3 - 2\sqrt{2})$, becomes uniformly distributed until $x = (3 - 2\sqrt{2})$, and then decreases to 0 as x increases to 1 . For $\alpha = 1$, the distribution is entirely concentrated at the point $x = 0$. From these results, it appears that the density function of the asymptotic distribution tends to concentrate towards the centre, from the value $\frac{1}{2}$ at $x = \pm 1$ when $\alpha = 0$ until finally the density function concentrates at the middle point $x = 0$ when $\alpha = 1$. Karlin, S. in [2] points out that for $0 \leq \alpha < \frac{1}{2}$, the distribution is on a Cantor-like set of points and conjectures that for $1 > \alpha \geq \frac{1}{2}$, the cumulative distribution becomes continuous.

Section Two

From chapter two equation (2.3.3) the moment generating function of the distribution of p-values for two subject-controlled events in the equal α case on the nth trial is given by the following functional differential equation

$$M_{n+1}(t) = e^{ta_2} M_n(\alpha t) + \frac{e^{ta_1} - e^{ta_2}}{\alpha} \frac{d}{dt} M_n(\alpha t)$$

and the asymptotic moment generating function satisfies (2.3.4)

$$M(t) = e^{ta_2} M(\alpha t) + \frac{e^{ta_1} - e^{ta_2}}{\alpha} \cdot \frac{d}{dt} M(\alpha t) \quad .$$

Attempts have been made to solve the above functional differential equations. However, so far, no explicit solution has been obtained. Only certain properties of the functional differential equations in the case $\alpha = \frac{1}{2}$, $\lambda_1 = 0$, $\lambda_2 = 1$ (i.e. $Q_1 p = \frac{1}{2} p$ and $Q_2 p = \frac{1}{2} p + \frac{1}{2}$) have been obtained and we consider these in the subsequent remarks.

(a) By substituting $\alpha = \frac{1}{2}$ into the moment generating functional differential equation (2.3.3) on the (n + 1)th trial

$$(3.2.1) \qquad M_{n+1}(t) = e^{t/2} M_n(t/2) + \frac{1 - e^{t/2}}{1/2} \frac{d}{dt} M_n(t/2) \quad .$$

Since as stated at the end of section three, chapter one, the asymptotic p-values in this case are independent of the initial p value , we may take $p_o = \frac{1}{2}$ for convenience.

Thus

$$(3.2.2) \qquad M_o(t) = e^{t/2}$$

$$M_1(t) = e^{t/2} M_o(t/2) + \frac{1 - e^{t/2}}{1/2} \frac{d}{dt} M_o(t/2)$$

$$= \frac{1}{2} e^{t/4} + \frac{1}{2} e^{t3/4}$$

$$M_2(t) = e^{t/2} M_1(t/2) + \frac{1 - e^{t/2}}{1/2} \frac{d}{dt} M_1(t/2)$$

$$= \frac{1}{8} e^{t/8} + \frac{3}{8} e^{t3/8} + \frac{3}{8} e^{t3/8} + \frac{1}{8} e^{t7/8}$$

and similarly

$$M_3(t) = e^{t/2} M_2(t/2) + \frac{1 - e^{t/2}}{1/2} \frac{d}{dt} M_2(t/2)$$

$$= \frac{1}{64} e^{t/16} + \frac{9}{64} e^{t3/16} + \frac{15}{64} e^{t5/16} + \frac{7}{64} e^{t7/16}$$

$$+ \frac{7}{64} e^{t9/16} + \frac{15}{64} e^{t11/16} + \frac{9}{64} e^{t13/16} + \frac{1}{64} e^{t15/16} .$$

Let the values of p on the nth trial be arranged in increasing order, and let $f_n(p_{r,n})$ be the probability of the rth value of p on the nth trial given by the coefficient of $e^{tp_{rn}}$ in the moment generating function. From the above, for any n ,

$$(3.2.3) \qquad M_n(t) = \sum_{r=1}^{2^n} f_n(p_{r,n}) e^{tp_{rn}}$$

where $\quad p_{r,n} = \dfrac{2r - 1}{2^{n+1}} , \qquad (r = 1, \dots 2^n) ,$

and

$$(3.2.4) \qquad M_{n+1}(t) = \sum_{r=1}^{2^{n+1}} f_{n+1}(p_{r,n+1}) e^{tp_{r,n+1}}$$

where $\quad p_{r,n+1} = \dfrac{2r - 1}{2^{n+2}} , \qquad (r = 1, \dots 2^{n+1}) .$

Using (3.2.1)

$$M_{n+1}(t) = e^{t/2} M_n(t/2) + 2(1 - e^{t/2}) \frac{d}{dt} M_n(t/2) ,$$

we obtain

$$(3.2.5) \quad M_{n+1}(t) = e^{t/2} \sum_{r=1}^{2^n} f_n(p_{r,n}) e^{\frac{t(2r-1)}{2^{n+2}}}$$

$$+ (1 - e^{t/2}) \sum_{r=1}^{2^n} f_n(p_{r,n}) \frac{(2r-1)}{2^{n+1}} e^{\frac{t(2r-1)}{2^{n+2}}}$$

$$= \sum_{r=1}^{2^n} p_{r,n} f_n(p_{r,n}) e^{tp_{r,n+1}}$$

$$+ \sum_{r=1}^{2^n} f_n(p_{r,n}) \frac{(2^{n+1} - 2r + 1)}{2^{n+1}} e^{\frac{t(2^{n+1} + 2r - 1)}{2^{n+2}}} .$$

In the second sum let $R = 2^n + r$. Then

$$(3.2.6) \quad M_{n+1}(t) = \sum_{r=1}^{2^n} p_{r,n} f_n(p_{r,n}) e^{tp_{r,n+1}}$$

$$+ \sum_{R=2^n+1}^{2^{n+1}} f_n(p_{R-2^n,n}) \frac{(2^{n+2} - 2R + 1)}{2^{n+1}} e^{\frac{t(2R-1)}{2^{n+2}}}$$

$$= \sum_{r=1}^{2^n} p_{r,n} f_n(p_{r,n}) e^{tp_{r,n+1}} + \sum_{r=2^n+1}^{2^{n+1}} f_n(p_{r-2^n,n}) p_{2^{n+1}-r+1,n} e^{tp_{r,n}}$$

Comparing (3.2.4) and (3.2.6) , we see that

$$(3.2.7) \quad f_{n+1}(p_{r,n+1}) = p_{r,n} f_n(p_{r,n}) \qquad \text{for } 1 \leq r \leq 2^n$$

$$(3.2.8) \quad f_{n+1}(p_{r,n+1}) = p_{2^{n+1}-r+1,n} f_n(p_{r-2^n,n}) \qquad \text{for } 2^n + 1 \leq r \leq 2^{n+1} .$$

From (3.2.7) and (3.2.8) , table 3.1 is constructed for $x = 0, 1, 2, 3, 4$. It is obvious from the table that the distribution is symmetrical for all values of n .

Table 3.1. Values of $f_n(p_{r,n})$ for $n = 0, 1, 2, 3, 4$. The values in each column of $f_n(p_{r,n})$ in the table divided by the corresponding $2^{n(n+1)/2}$ give the probabilities of the p values on the nth trial. As this is a two subject-controlled event, there are 2^n possible p values on the nth trial.

r	n / numerator of p_r	$f_n(p_{r,n})$ 0	1	2	3	4
	$2^{n(n+1)/2}$	1	2	8	64	1024
1	1	1	1	1	1	1
2	3		1	3	9	27
3	5			3	15	75
4	7			1	7	49
5	9				7	63
6	11				15	165
7	13				9	117
8	15				1	15
9	17					15
10	19					117
11	21					165
12	23					63
13	25					49
14	27					75
15	29					27
16	31					1

It can be seen from the table that the (2^{n-1})th and $(2^{n-1}+1)$th p values on the nth trial always have a probability of $\frac{2(2^{n-1})-1}{2^{n(n+1)/2}}$. Since there are 2^n p-values on the nth trial and they are equally spaced, we can consider the unit interval as divided into 2^n equal subintervals. For example for $n = 3$, $2^n = 8$ and this is shown in figure 3.9.

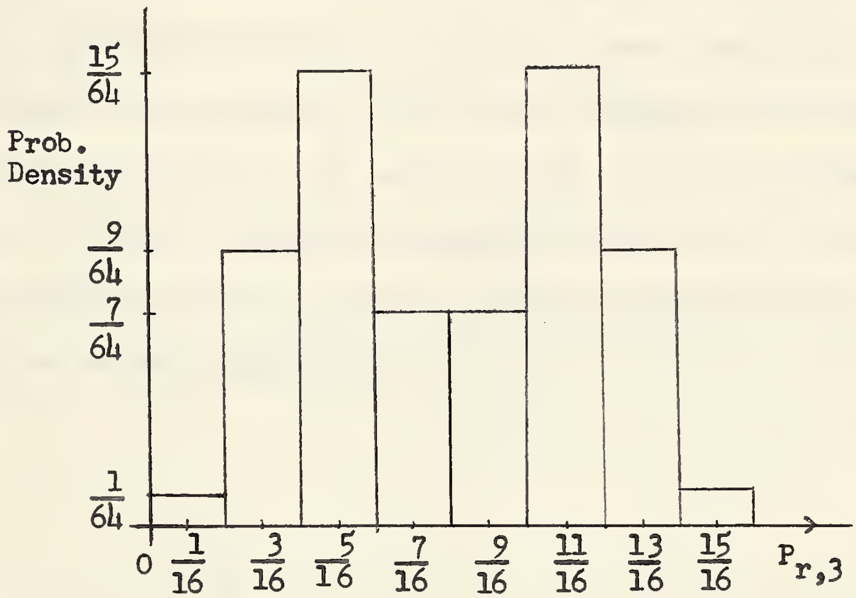

Fig. 3.9. Histogram of $P_{r,3}$ where $r = 1, 2, 3, 4, 5, 6, 7, 8$.

The width of each subinterval is $\frac{1}{2^n}$. Since

Probability = probability density x width of interval,

we have for $r = 2^{n-1}$

$$(3.2.9) \quad \frac{2(2^{n-1})-1}{2^{n(n+1)/2}} = \text{prob. density} \times \frac{1}{2^n}$$

or

$$(3.2.10) \quad \text{Prob. density of } p_{2^{n-1},n} = \frac{2^n(2^n-1)}{2^{n(n+1)/2}}$$

$$\leq \frac{2^{2n}}{2^{n(n+1)/2}} \quad .$$

Thus as n approaches infinity,

$$(3.2.11) \quad \lim_{n \to \infty} p_{2^{n-1},n} < \lim_{n \to \infty} \frac{2^{2n}}{2^{n(n+1)/2}} \longrightarrow 0 \quad .$$

Similarly the probability density of $p_{2^{n-1}+1,n}$ also approaches 0 as n approaches infinity. Thus in the asymptotic distribution the probability density of $p = \frac{1}{2}$ is zero. The table itself suggests the conjecture that in the asymptotic distribution the maximum lies at $p = \frac{1}{3}$ and at $p = \frac{2}{3}$. Dr. T. V. Narayana [10] has devised a convincing argument that this is so. From the above results, it is likely that the asymptotic distribution is of the shape shown in figure 3.10.

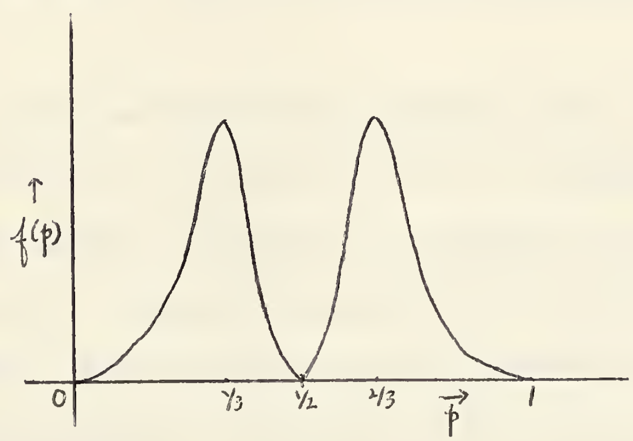

Fig. 3.10. The conjecture of the asymptotic distribution of two subject-controlled events with $\alpha_1 = \alpha_2 = \frac{1}{2}$ and $\lambda_1 = 0$ and $\lambda_2 = 1$.

The shape of the curve of figure 3.10 is in close agreement with the result obtained by R. R. Bush and F. Mosteller in [1] page 138,

with a 2000-trial stat rat. Fig. 3.11 is a reproduction of their work.

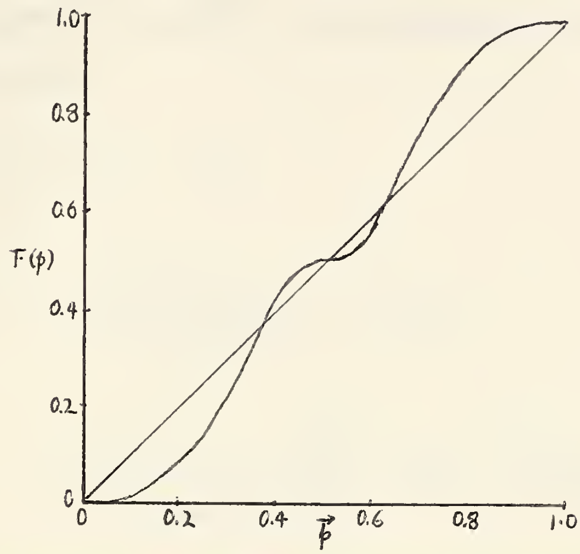

Fig. 3.11. The approximate asymptotic cumulative distribution of p values, obtained from a 2000 trial stat rat, for the case of two subject-controlled event with $Q_1 p = 0.5p$ and $Q_2 p = 0.5 + 0.5p$. The straight line corresponds to a uniform distribution from zero to unity. (Reproduced from [1] page 138)

From figure 3.11, the approximate cumulative curve appears to be symmetrical about $p = \frac{1}{2}$ and at $p = \frac{1}{2}$, it appears to have a slope equal to zero. Furthermore, it appears that at $p = \frac{1}{3}$ and at $p = \frac{2}{3}$, the cumulative curve is at a point of inflexion respectively. Thus it appears that the probability density distribution curve has a maximum at $p = \frac{1}{3}$ and another one at $p = \frac{2}{3}$ respectively. This is in agreement with the result obtained above from the consideration of the moment generating functional equation. From equation (1.3.7) the mean of the asymptotic distribution is

$$(3.2.12)\quad V_{1,\infty} = \frac{a_2}{1 - a_1 + a_2 - \alpha} = \frac{1 - \alpha}{1 - (2\alpha - 1)} = \frac{1}{2}$$

and from equation (1.3.15), the variance of the asymptotic distribution is

$$(3.2.13)\quad \sigma_\infty^2 = V_{2,\infty} - V_{1,\infty}^2$$

$$= \frac{a_2{}^2 + B_1 V_{1,\infty}}{1 - B_2} - V_{1,\infty}^2$$

$$= \frac{1}{1 - \alpha(3\alpha - 2)} \left[(1 - \alpha)^2 + \frac{(1 - \alpha)^2(3\alpha - 1)}{1 - (2\alpha - 1)} \right] - (\tfrac{1}{2})^2$$

$$= \frac{1 + \alpha}{2(1 + 3\alpha)} - (\tfrac{1}{2})^2$$

$$= \frac{1 - \frac{1}{2}}{4(1 + 3/2)}$$

$$= 0.05 \ ,$$

so that $\sigma_\infty = 0.2236$.

The values obtained from the single stat-rat computation were $V_{1,\infty} = 0.4958$ and $\sigma_\infty = 0.2227$.

(b) The characteristic function of the asymptotic distribution is from (2.3.9) given by the equation

$$\phi(t) = e^{ita_2}\,\phi(t/2) - 2i\,(e^{ita_1} - e^{ita_2})\,\frac{d}{dt}\,\phi(t/2) \quad .$$

Since in this case $\lambda_1 = 0$, $\lambda_2 = 1$, we have $a_1 = 0$ and $a_2 = \frac{1}{2}$. Substituting $a_1 = 0$, $a_2 = \frac{1}{2}$ and t by $2t$, the characteristic function becomes

$$\phi(2t) = e^{it}\,\phi(t) - i\,(1 - e^{it})\,\frac{d}{dt}\,\phi(t) \quad , \quad \text{or}$$

$$\frac{d}{dt}\,\phi(t) + \frac{ie^{it}\phi(t)}{1 - e^{it}} = i\,\frac{\phi(2t)}{1 - e^{it}} \quad . \tag{3.2.14}$$

To transform the asymptotic p-distribution interval from $[0,1]$ to $[-1,1]$ let $x = 2p - 1$ where $-1 \leq x \leq 1$. As in (3.1.34) $\phi_x(t)$, the characteristic function of the asymptotic distribution of x is equal to

$$\phi_x(t) = e^{-it}\,\phi(2t) \quad , \quad \text{or}$$

$$\phi(t) = e^{it/2}\,\phi_x(t/2) \quad .$$

Thus the differential equation satisfied by $\phi_x(t)$ is from (3.2.14)

$$2(1 - e^{it})\,\frac{d}{dt}\,\phi_x(t/2) + i(1 + e^{it})\,\phi_x(t/2) = 2ie^{it/2}\,\phi_x(t) \quad . \tag{3.2.15}$$

For convenience let $t = 2\theta$, then (3.2.15) becomes

$$(3.2.16) \quad (1 - e^{2i\theta}) \frac{d}{d\theta} \phi_x(\theta) + i(1 + e^{2i\theta}) \phi_x(\theta) = 2ie^{i\theta} \phi_x(2\theta)$$

$$\frac{1}{i}(e^{-i\theta} - e^{i\theta}) \frac{d}{d\theta} \phi_x(\theta) + (e^{-i\theta} + e^{i\theta}) \phi_x(\theta) = 2\phi_x(2\theta) \text{ , or}$$

$$(3.2.17) \quad \cos\theta \cdot \phi_x(\theta) - \sin\theta \frac{d}{d\theta} \phi_x(\theta) = \phi_x(2\theta) \text{ .}$$

An attempt was made to solve this functional differential equation by the Fourier series. However, the odd coefficients of the series are obtained only in the form of a recurrence equation which I have not been able to solve.

From remark (1), it has been shown that the distribution is symmetrical about the centre and so $\phi_x(\theta)$ must be an even function. Thus let

$$(3.2.18) \quad \phi_x(\theta) = \frac{a_o}{2} + \sum_{r=1}^{\infty} a_n \cos n\theta \text{ ,}$$

where $\frac{a_o}{2}$ and a_n $(n = 1, 2, \ldots \infty)$ are constants of the Fourier cosine series. Substituting this into the differential equation (3.2.17),

$$(3.2.19) \quad \cos\theta \left[\frac{a_o}{2} + \sum_{n=1}^{\infty} a_n \cos n\theta\right] + \sin\theta \sum_{n=1}^{\infty} a_n n \sin n\theta$$

$$= \frac{a_o}{2} + \sum_{n=1}^{\infty} a_n \cos 2n\theta \text{ .}$$

This functional differential equation is equal to

$$\frac{a_o}{2}\cos\theta + \sum_{n=1}^{\infty} a_n \cos\theta\cos n\theta + \sum_{n=1}^{\infty} n\, a_n \sin\theta \sin n\theta$$

$$= \frac{a_o}{2} + \sum_{n=1}^{\infty} a_n \cos 2n\theta \,, \quad \text{or}$$

$$\text{(3.2.20)} \quad \frac{a_o}{2}\cos\theta + \sum_{n=1}^{\infty} a_n \tfrac{1}{2}\left[\cos(\theta + n\theta) + \cos(\theta - n\theta)\right]$$

$$+ \sum_{n=1}^{\infty} na_n \tfrac{1}{2}\left[\cos(\theta - n\theta) - \cos(\theta + n\theta)\right]$$

$$= \frac{a_o}{2} + \sum_{n=1}^{\infty} a_n \cos 2n\theta \;.$$

Consider the even coefficients. Let $n = 2k$ where $k = 1, 2, 3, \ldots$. Equating the coefficients of the cosines

$$\text{(3.2.21)} \quad a_{2k}\, \tfrac{1}{2}(1 - 2k)\cos(2k + 1)\theta + a_{2k+2}\, \tfrac{1}{2}(1 + 2k + 2)\cos(2k + 1)\theta = 0$$

i.e.

$$\text{(3.2.22)} \quad a_{2k}(1 - 2k) + a_{2k+2}(2k + 3) = 0 \;, \quad \text{or}$$

$$a_{2k+2} = \frac{2k - 1}{2k + 3}\, a_{2k} \;.$$

Thus by (3.2.22)

(3.2.23) $$a_{2k} = \frac{2k - 3}{2k + 1} a_{2k-2}$$

$$= \left(\frac{2k - 3}{2k + 1}\right)\left(\frac{2k - 5}{2k - 1}\right) \cdots \left(\frac{3}{7}\right)\cdot\left(\frac{1}{5}\right)\cdot\left(\frac{-1}{3}\right)\cdot a_o$$

$$= \frac{-a_o}{(2k + 1)(2k - 1)} \quad .$$

Equating the coefficients of $\cos 2n\theta$ of (3.2.20)

$$a_{2n-1}(1 - n) \cos 2n\theta + a_{2n+1}(n + 1) \cos 2n\theta = a_n \cos 2n\theta \quad \text{or}$$

(3.2.24) $$a_n = a_{2n-1}(1 - n) + a_{2n+1}(1 + n) \quad .$$

Thus the odd coefficients are given by the recurrence relation

(3.2.25) $$a_{2n+1} = \frac{a_n - a_{2n-1}(1 - n)}{1 + n} \quad .$$

Attempts have been made without success to solve this recurrence equation.

(c) By (3.2.17) the functional differential equation satisfied by the characteristic function of the asymptotic distribution of $x = 2p - 1$ is

$$\phi_x(2\theta) = \cos\theta\, \phi_x(\theta) - \sin\theta \frac{d}{d\theta} \phi_x(\theta) \quad .$$

This functional differential equation may be simplified by the substitution

$$(3.2.26) \quad \phi_x(\theta) = \sin\theta \, g_x(\theta) \ ,$$

so that

$$(3.2.27) \quad \frac{d\,\phi_x(\theta)}{d\theta} = \cos\theta \, g_x(\theta) + \sin\theta \, \frac{d\,g_x(\theta)}{d\theta}$$

yielding

$$(3.2.28) \quad \frac{d\,g_x(\theta)}{d\theta} = -2\cot\theta \, g_x(2\theta) \quad .$$

Attempts have been made to solve the functional differential equation in this form. However, they have not been successful.

BIBLIOGRAPHY

[1] Bush, Robert R. and Mosteller, Frederick. Stochastic models for learning. New York: John Wiley and Sons Inc., 1955.

[2] Karlin, S. "Some random walks arising in learning models 1." Pacific Journal of Mathematics, 3, (1953), 725 - 756.

[3] Harris, T. E. "Personal communication and abstract." Annals of Mathematical Statistics, 23, (1952), 141.

[4] Kendall, M. G. The advanced theory of statistics, Volume I. London: C. Griffin and Company, 1943.

[5] Kac, Mark. Statistical independence in probability analysis and number theory. New York: Carus Mathematical Monograph. The Mathematical Association of America, 1959.

[6] Peirce, B. O. A short table of integrals. Boston: Ginn and Co., 1929.

[7] Erdelyi, A. et al. Bateman Manuscript Project. California Institute of Technology. Table of integral transforms. New York: McGraw-Hill, 1954.

[8] Titchmarsh, E. C. The theory of functions. Oxford University Press, 1939.

[9] Feller, W. An Introduction to Probability Theory and Its Applications. New York: John Wiley and Sons Inc., 1957.

[10] Narayana, T. V. Personal communication.